ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

VOLUME 5

1937

ECONOMETRICA

Published at 450 Ahnaip St., Menasha, Wis.

BUSINESS OFFICE OF ECONOMETRICA, Mining Exchange Building, Colorado Springs, U. S. A. In the absence of specifications to the contrary, any communication regarding ECONOMETRICA or the Econometric Society, may be addressed to this office.

The Econometric Society is not responsible for theories and opinions published in Econometrica. The authors assume sole responsibility.

Published quarterly for the Econometric Society by the George Banta Publishing Company, 450 Ahnaip St., Menasha, Wisconsin.

Subscription Price $7.00 per annum. Single copies $1.75.

INDEX

ARTICLES

MEETINGS

MISCELLANEOUS

INCOME IN THEORY AND INCOME TAXATION IN PRACTICE*

By Irving Fisher

CHAPTER I: INTRODUCTION

Elsewhere[1] I have defined the income of a person as consisting of services rendered to him. These services may come to him by virtue of his property (including his claims to, or in, physical wealth and his claims on other people for service), or they may come to him as services rendered by himself—that is, by work.

Services are also classified into (1) services in kind, such as the use of one's own dwelling and (2) the bringing in of money, such as interest from bonds or dividends from stocks.

For the purposes of this study we may best begin by considering only the latter kind of services; that is, services in the form of money receipts, whether credited to one's own labor or to physical wealth such as real estate or to property rights in real wealth such as stocks or bonds. Services in kind we shall at first ignore. After we have gone as far as possible in the analysis of money receipts, services in kind will be introduced to complete the picture.

Originally "income" was probably thought of as simply *incoming money*. Incoming payments of money naturally appeared in sharp contrast with outgoing payments of money. A business man, in his shop, could easily subtract the outgoings from the incomings, call what was left his "net income," and physically take this "net income" out of his business shop into his home. It was called net income because it was the net money coming in to his home from his business. But today such simple accounting has been superseded, or overlaid, by many complicated procedures; modern accountancy has evolved into an elaborate art. It has done so almost without help from economists. The result is that the chasm between accountancy and economics has become wide.

So far as I know, only two serious attempts have been made to bridge this chasm, one being my *Nature of Capital and Income*, first published in 1906 and the other, *The Economics of Accountancy* by Professor John B. Canning, published in 1929. My book started from the econo-

* This paper was presented in preliminary form in four lectures on July 7 to 10, 1936 at the Cowles Commission Research Conference on Economics and Statistics at Colorado Springs.

[1] *The Nature of Capital and Income*, Macmillan, 1906, which see for numerous associated definitions and concepts. These are all assembled at the end in a "glossary." See also "Are Savings Income?" *American Economic Association Quarterly*, Third Series, Vol. 9, No. 1, April, 1908, pp. 21–22.

mists' side of the chasm; Canning's, from the accountants' side. Together they bridge the gap. They also meet, midstream, and fit into each other well.

When *The Nature of Capital and Income* appeared, there was so much less interest in mathematical economics than today that I endeavored to make the exposition, as far as possible, verbal, supplemented by diagrams, with a minimum of mathematics relegated to Appendices. But when a mathematical subject is discussed in non-mathematical language, there is danger of misunderstanding.

One contention in particular has been a stumbling block. This is that neither the undistributed profits of a corporation nor, in fact, any gains in capital value are properly income. They are capital; and to tax them as income and later to tax the income flowing from them is a form of double taxation. This contention that capital gain, appreciation, undivided profits, savings, are, in the last analysis, not properly income has been sometimes dubbed "Fisher's *pons asinorum*."[2]

As I have said elsewhere:

> Income plays an important role in all economic problems; it is income for which capital exists; it is income for which labor is exerted; and it is the distribution of income which constitutes the disparity between rich and poor.[3]

Paraphrasing another statement made by me in *The American Economic Review*, Professor John B. Canning has said:

> *Income is, without exception, the simplest and most fundamental concept of economic science; only by means of this concept can other economic concepts ever be fully developed and understood, and upon beginning with this concept depends the full fruition of economic theory in economic statistics.*[4]

With the increasing use of mathematics by economic students, the concept of income which best lends itself to mathematical analysis may be expected to be accepted more generally. Thereupon four great interests ought to leap forward, namely economic theory, statistics,[5] accountancy, and property law,[6] and especially income-tax law; for these

[2] See Professor A. W. Flux, "Irving Fisher on Capital and Interest," *Quarterly Journal of Economics*, February, 1909. That capital gain is not properly to be called income was first enunciated by me in December, 1897 in "The Role of Capital in Economic Theory," *Economic Journal*, pp. 532–3.

[3] *The Nature of Capital and Income*, Preface.

[4] John B. Canning, Ph.D., *The Economics of Accountancy*, The Ronald Press, 1929, p. 175.

[5] See Morris A. Copeland, "National Wealth and Income—An Interpretation," *Journal of the American Statistical Association*, June, 1935, p. 377.

[6] See Edwin A. Howes, Jr., *The American Law Relating to Income and Principal*, Little, Brown, and Company, 1905, 104 pp.; Walter Strachan, *A Digest of the Law of Trust Accounts*, Effingham Wilson, 1911, 186 pp.; J. M. Maguire, "Capitalization of Periodical Payments by Gift," *Harvard Law Record*, Novem-

four interests will be released from important confusions now holding them back.

In this monograph the effort has been made, as far as possible, not to repeat *The Nature of Capital and Income*, but to supplement it both by new methods and by new results. Some of the new material is pure economics—almost pure mathematics. And—what should be especially emphasized—it is this pure mathematical economics on which we should largely rely for the conclusive proof that much of our present income-tax legislation is confused in theory and harmful in practice.

One aim of this study is to help frame a plan for an income tax which can be easily applied in practice while being, at the same time, sound in theory.

What seems especially to call for a new study is the adoption in 1936, in the United States, of a new tax on undivided profits. Whatever may be said in its behalf, it may be demonstrated to involve double taxation in so far as undivided profits are actually taxed.

The very fact that, for so many years, we have permitted capital levies, such as the tax on capital gain, to masquerade as income taxes is eloquent testimony to the need for a better understanding of this troubled subject.

That something is wrong with our capital-gain tax has long been apparent. Its absurd results have enabled rich men lawfully to avoid income taxes, and has made such tax-avoidance an object of popular ridicule. Anyone can see the absurdity of J. P. Morgan and his partners proving by the present laws that their incomes were less than nothing, and of Raskob and DuPont registering capital losses by selling each other stocks as well as the like absurdities in the cases of Mellon and Mitchell. But what the public fails to see is that the absurdity reflects not on those men but on the tax laws which, apparently, they fully observed.

When *The Nature of Capital and Income* was written it was intended as pure economic theory and was written largely to serve as a foundation for a later book, *The Rate of Interest*. There was then no income tax in the United States; hence the book contained no practical directions as to how a sound income tax should be formulated. When, later, others attempted to draw from that book a practical scheme for an income tax, they concluded that the thing was impossible, the gap between theory and practice being too wide. Even in my subsequent writings, in which some of the absurdities of the income tax were pointed out, the practical task of reforming our tax laws was left for others.

ber, 1920, pp. 20–49; Walter Strachan, "A Company's Capital or Income," *The Law Quarterly Review*, July, 1930, pp. 334–340; Roswell Magill, *Taxable Income*, Ronald Press, 1936.

Now that I have myself sought the solution of this practical problem, in the light of the theory outlined in *The Nature of Capital and Income,* it turns out to be very simple—far simpler than our present income tax; simpler even than England's income tax, though that country long ago cast out most of the impossible capital-gain tax still plaguing America.

But the contribution of this monograph to the solution of the great problem of scientific taxation will be restricted exclusively to showing how income may be most usefully defined. There will be no discussion of whether an income tax should be progressive, much less of what place an income tax should have in a complete scheme of taxation.

CHAPTER II: FORMULAE ON MONEY RECEIVED FROM INVESTMENTS

Before discussing what changes in our income-tax laws would follow from correcting the legislative definition of income, it is necessary to discuss the foregoing concept of income. We shall discuss first money-receipts from property (investments); then money-receipts from labor; and finally services in kind.

To fix our ideas, let us take an individual, John Smith, and designate by the letter M any receipt by him, that is, payment to him, of money. Such payments are *incoming* payments; but outgoing payments may, for the present, be included as M's having a negative value.

It will make our analysis seem more definite and realistic if we thus confine the discussion, for the present, to money-receipts. Also, until further notice, we shall assume, as in the case of a safe bond,

(1) that each payment, M, is of a *definitely specified* sum of money;

(2) that it is made, or rather is to be made, at a *definite and foreknown time,* and

(3) that there is no doubt whatever about it—i.e., *no risk* of default.

To be as general as possible, within the limits just prescribed, let the time elapsing up to any payment be measured from any origin called "the present" and let such elapsed *t*ime be denoted by t, where t may be 1, 2, 3, etc., units of time.

The unit of time chosen may, of course, be of any length. Theoretically it will make a difference to the discounted value of a payment whether the payment occurs at the beginning, end, middle, or elsewhere within this time unit. But by taking the time unit sufficiently small—reducing it, if desired, even to a day or a second—we may make any such difference negligible. Therefore, no impairment of generality will occur if, for simplicity, we assume, as we shall, that each payment takes place at the *end* of the given time unit. In practice, few contracts are drawn for two or more payments closer together than three months —a "quarter." The most convenient unit for us to use, however, is the

year. Therefore let us think hereafter in terms of this unit, the year, though the same principles would apply to any other time unit.

Let M_t denote the number of *m*oney units (say dollars) received by John Smith in any given unit of *t*ime (between $t-1$ and t), said M_t being located at the end of that time unit.

Let the rate of *i*nterest be i, assumed for simplicity to be the same for all years and to be calculated as if payable annually.[7]

This one magnitude i and the series of magnitudes M_t being given, all the other magnitudes with which we are here concerned can be derived. The first to be derived is the capitalized, discounted, or present, value of the future M's.

Let M_1 be a sum of money receivable one year from date. Let its present value (at time $t=0$) be C_0. Then the rate of interest, i, is defined by the equation $1+i=M_1/C_0$ or $C_0=M_1v$, where[8] v is $1/(1+i)$.

Similarly, if C_0 is exchanged for M_2, two years in the future, the i realized is such that $C_0=M_2v^2$; and, in general $C_0=M_qv^q$ where q is any given value of t.

If C_0 denotes the discounted value of a *series* of M's located at times $t=1, 2, \cdots n$, then

$$C_0 = \sum_{t=1}^{t=n} M_t v^t.$$

To be more general, instead of taking the present value simply at time $t=0$, we may take it at any time $t=q$ ($q=0, 1, \cdots n$). The above formula then becomes

$$C_q = \sum_{t=q+1}^{t=n} M_t v^{t-q}. \tag{1}$$

[7] The equivalent rate of interest per annum, payable *semi*-annually, is i', such that $1+i=(1+i'/2)^2$. The equivalent rate per annum payable *quarterly* is i'', such that $1+i=(1+i''/4)^4$. The equivalent rate per annum payable *momently* is $i'''^{\cdots}$, such that $1+i=\lim\ (1+i'''^{\cdots}/n)^n$ where n increases indefinitely. The last expression may be written $\lim\ [(1+i'''^{\cdots}/n)^{n/i'''^{\cdots}}]^{i'''^{\cdots}}$ or $e^{i'''^{\cdots}}$ where e is the base of the Naperian, or "natural," system of logarithms (approximately 2.718).

It is here assumed, for simplicity, that i is the same for all the M's discounted and that, in a series of M's, the same i appiies to each M individually as applies to them all, or to any number of them as a group.

[8] Since $1/(1+i)=v$, it follows that $1-v=i/(1+i)$, and $(1-v)/v=i$. These forms will be used later.

Following common usage, all discounting is here expressed in terms of i, the rate of *interest*. It might seem more appropriate to express everything in terms of a rate of *discount*, d, such that $v=1-d=1/(1+i)$ so that $(1-d)(1+i)=1$. The magnitude i always slightly exceeds the magnitude d but the two magnitudes become equal at the limit if the time interval be decreased indefinitely. Then i becomes successively i, i', i'', and, at the limit, $i'''^{\cdots}$, and d becomes successively d, d', d'', and, at the limit, $d'''^{\cdots}$.

Equation (1) is the general[9] formula for the capital value, C_q, of a specified series of M's, under the present three hypotheses.

We may define capital gain (for any year ending at time $t=q$ immediately after M_q has been received) as

$$G_q = C_q - C_{q-1}. \tag{2}$$

Capital gain, as above defined, does not imply that it has been "realized" in cash. It simply means that the gain has "accrued." Moreover, this accrual is solely due to the drawing near of the future receipts, the M's, of which the C's are the discounted values.

The "earnings" of capital (denoted by E_t for year ending at time t) may be defined as

$$E_q = M_q + G_q. \tag{3}$$

Evidently earnings are equal to receipts if capital gain is zero, and are equal to capital gain if receipts are zero. The former case is exemplified by a bond which remains at par; the latter case, by a promissory note, or savings-bank account, drawing compound interest.

Besides equations (1), (2), (3), there are numerous other equations connecting the M's, C's, G's, and E's.

For instance, we may obtain an E in terms of a C as follows:

We know by (1),

$$C_0 = \sum_{t=1}^{t=n} M_t v^t = M_1 v + \sum_{t=2}^{t=n} M_t v^t.$$

Also,

$$C_1 = \sum_{t=2}^{t=n} M_t v^{t-1}.$$

[9] Incidentally, to complete the picture of C in all its values between $t=0$ and $t=n$, we may add that, immediately after $t=n$, there being no further receipts, capital value becomes zero, that is $C_{n+\delta}=0$, where $\delta>0$.

Moreover, capital value for $t=0-\delta$ will also be zero if we suppose that, at time $t=0$, there is an item M_0 of negative value, equal, though opposite, to C_0 (which is the capital value *immediately after* said negative item M_0 is paid out). This M_0 would denote the original payment *out* (or investment made) for the right to the future series of payments *in*, $M_1, M_2 \cdots, M_n$.

Thus, under the conditions stated, the series of capital values C_0, C_1, $C_2, \cdots, C_n$, is theoretically preceded by, and also followed by, a zero capital value. Like life, C goes "from dust to dust." Between its birth and death, capital rises in anticipation of the M's ahead of it and, as each (positive) M is passed, falls by the amount of that M.

Or if, reversely, we trace its value backward, C is zero just *after* M_n, and is M_n just *before* said M_n; it is $M_n v$ just after M_{n-1} and is $M_n v+M_{n-1}$ just before said M_{n-1}; and so on back, step by step, fulfilling in all cases equation (1). The figure at the end of this chapter illustrates this.

Multiplying the last equation by v, and subtracting this from the first above, we obtain

$$C_0 - vC_1 = M_1 v, \quad \text{or},$$

$$C_0 = v(C_1 + M_1) = v(C_0 + G_1 + M_1) = v(C_0 + E_1).$$

Whence

$$E_1 = C_0 \frac{(1 - v)}{v},$$

which, in accordance with footnote 8, reduces to $C_0 i$.

By the same reasoning, more generally,

$$E_q = C_{q-1} i. \tag{4}$$

Thus E_q, besides being equal, by definition, to $M_q + G_q$, is also equal to $C_{q-1}i$; that is, E_q is the "interest on the capital"; that is, the capital value at any time equals earnings in the ensuing year divided by the rate of interest (or multiplied by its reciprocal, the "price-earnings" ratio).

Thus we may (under our assumptions) derive any C from current earnings without knowing the earnings of future years. This is quite unlike the dependence of C on the M's.

We may, if we wish, start with equation (4) as our definition of E and, reversing the above reasoning, derive equation (3).

Many other simple relations among the four sets of magnitudes may be derived, such as:

$$C_q = -\sum_{t=q+1}^{t=n} G_t; \tag{5}$$

$$E_q = -\sum_{t=q}^{t=n} G_t i; \tag{6}$$

$$G_q = (E_{q+1} - E_q) \div i; \tag{7}$$

$$M_q = C_{q-1}(1 + i) - C_q; \tag{8}$$

$$M_q = [E_q(1 + i) - E_{q+1}] \div i; \tag{9}$$

$$M_q = -\sum_{t=q}^{t=n} G_t i - G_q; \tag{10}$$

$$E_q = \sum_{t=q}^{t=n} M_t v^{t-q}(1 - v); \tag{11}$$

$$G_q = \sum_{t=q+1}^{t=n} M_t v^{t-q}(1 - v) - M_q v; \tag{12}$$

$$C_0 = \sum_{t=1}^{t=q} M_t v^t + C_q v^q; \tag{13}$$

$$C_0 = \sum_{t=1}^{t=q} (M_t v^t + G_t v^q) + C_0 v^q. \tag{14}$$

Each of these equations contains only two sorts of magnitudes except formula (14) which contains three sorts, as does formula (3). Formulae (13) and (14) are part-way formulae, in which the summation extends, not to the nth year, but only to the qth year, leaving, in each case, a remainder term to take care of the intervening items. None of the 14 formulae[10] so far given, beyond the first four, will be needed in this study except equation (14) which shows the great difference between the roles played by the M's and the E's. That is, in general, C_0 is *not* $\sum_{t=1}^{t=n} E_t v^t$. Nor does this last magnitude, the capitalized series of earnings, play any role in our study.

Thus, equation (14) tells us that C_0 is the discounted value of three items:

(1) the M's, up to time q, each discounted from the future time at which it will actually occur.

(2) the G's, up to the time q, each *displaced* to time $t=q$ and then all discounted uniformly from there.

(3) C_0 itself, likewise displaced to time $t=q$ and then discounted from there.

It will thus be seen that one reason why C_0 is not the capitalized value of $E_1, E_2, E_3, \cdots E_n$, each E being the sum of an M and a G is that, while the M in each such sum is discounted backward from the time, t, at which it occurs, the G component in that sum is not so discounted, but must first be *displaced* forward to time q.

Yet it is often carelessly stated that capital value is the present value of future expected earnings.

Figure 1 below[11] illustrates both equation (1) and equation (14).

[10] The reader who so wishes should have no difficulty in verifying these equations. One way to verify them is to substitute for each magnitude its value as given in the equations of definition. The following is a skeleton of the simplest way to derive these equations. Equation (5) is found from equation (2) by repeating it for $t=q+1$, $t=q+2$, etc., and then adding these all together, remembering that $C_n=0$. Equation (6) is derived from (4) and substituting for C_{q-1} by (5). (7) comes from (2) substituting for the C's by (4). (8) is derived analogously to (4). (9) comes from (8) substituting by (4). (10) comes from (3) substituting by (6). (11), from (4) by (1). (12), from (2) by (1). (13), from (1) written for $t=0$ and for $t=q$ and subtracting. (14), from (13) and substituting $C_q=C_0+(C_q-C_0)=C_0\sum_{t=1}^{t=q}G_t$.

[11] For extended discussion of substantially the same problem, with the aid of charts, see "Are Savings Income?" already cited, especially pp. 30, 45, 46.

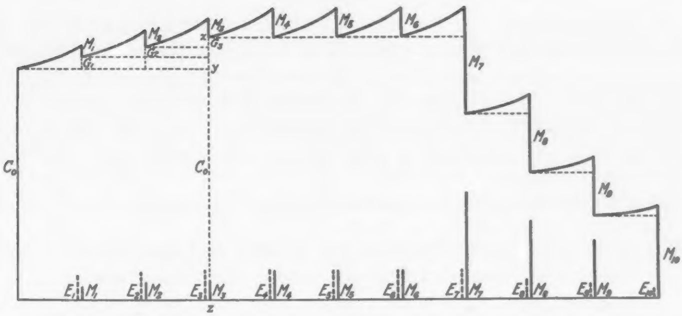


FIGURE 1

This figure shows the behavior of C, G, and E as derived from: (1) the ten given M's (each represented by a vertical line drawn upward from the base line); and (2) the rate of interest, i, on which the slope of the ten discount curves depends. The curve representing C is constructed backward, beginning with M_{10} (at time $t=10$ years) as follows: From the top of M_{10} (at the extreme right) draw (leftward) a discount curve until it reaches one year earlier (time $t=9$). Then add (upward), M_9. Then draw a discount curve (leftward) one year. Then add M_8; and so on, until we reach the "present," ten years before M_{10}, when the height is C_0.

If now, we reverse, and trace the curve forward in time, we see that, the first year, the curve gradually ascends, rising in all by an amount C_0i or G_1+M_1. It then suddenly falls by M_1, leaving a net gain of G_1. Then it rises by G_2+M_2 and falls back by M_2; etc., until, finally, it sinks to zero after the tenth and last payment, M_{10}.

The respective rises, G_1+M_1, G_2+M_2, etc., are the earnings, E_1, E_2, etc. These are also represented by dotted lines drawn from the base line, alongside of the M's. Compare now the E's and M's. Evidently the E's are the larger for the first three years; the E's and M's are equal for the next three years (so that the G's are zero); and the E's are the smaller for the last four years (so that the G's are negative).

C_0 is the discounted value of the ten M's. If desired, all ten of the discount curves may be continued leftward to "the present," dividing C_0 there into ten components respectively equal to the discounted values of the ten respective M's. *But C_0 is not the discounted value of the ten E's*, unless the G portions of these E's be displaced as described above.

Formula (14) is illustrated by drawing the vertical at time $t=3$. C_0, at "the present," is now seen as the discounted value of three groups of magnitudes, namely (1) M_1, M_2, and M_3, each being discounted from the time it occurs; (2) $G_1+G_2+G_3$, all together discounted from $t=3$ (constituting there the line xy; (3) C_0 also displaced to the time $t=3$ (constituting there the line yz). Evidently components (2) and (3), i.e., line xz, is a remainder term, being the discounted value, at time $t=3$, of all the M's remaining thereafter, namely M_4, M_5, $\cdots$ M_{10}.

If instead of time $t=3$, some other time is taken, the G's and C_0 must be further displaced to correspond; but the M's are always discounted without displacement.

Several of the other formulae may likewise be illustrated by Figure 1.

If, in accordance with footnote 9, we insert in the figure an M_0 equal to C_0 but opposite (drawn downward) to signify the original investment, we may illustrate an interesting corollary of equation (3). By (3), for every time unit, $E_t = M_t + G_t$. It follows that, for any time interval, $\sum E_t = \sum M_t + \sum G_t$. This equation applied to the complete time interval between t_{0-}, when C_{0-} is zero, and t_{n+}, when C_n is also zero (for which complete interval, therefore, $\sum G_q = 0$), gives $\sum_{t=0-}^{t=n+} E_t = \sum_{t=0-}^{t=n+} M_t$. That is, for the complete interval, the earnings and the monies out and in, are simply different distributions of the same total sum.

The same proposition holds true for any other interval for which there is no net capital gain. Perhaps an intuition of this conditional equality of an E sum and an M sum underlies the unfortunate confusion of E and M as "income."

The above proposition has been stressed in the writings of Professor John B. Canning.[12]

CHAPTER III: SPECIAL CASES

Equations (1)–(14) cover the general theory of money receipts from property under the three assumptions stated in Chapter II.

These 14 general relations include, of course, innumerable special cases. Among these, Perpetual Annuities and Bonds deserve special attention here, because they will help later in discussing certain Norms of accounting.

The perpetual annuity is an infinite series of equal payments at equal intervals, that is, $M_t = k$, a constant for all values of t from unity to infinity. The C's will also be constant. For the formula for C_0 is the same infinite series,

$$\sum_{t=1}^{t=\infty} kv^t,$$

as for C_1, C_2, etc.,

Therefore each year's capital gain is zero.

Each year's earnings are k. For by formula (3), $E_q = M_q + G_q = M_q = k$.

By formula (4), $C_q = E_{q+1}/i = k/i$; or $i = k/C_q$.

It should be remembered that i (in $v = 1/(1+i)$) was originally assumed to be invariable. Otherwise C_0 would not necessarily be equal

[12] *Economics of Accountancy*, pp. 94–96; pp. 126–127; also "A Certain Erratic Tendency in Accountants' Income Procedure," ECONOMETRICA, Vol. 1, No. 1, Jan., 1933, pp. 52–62. Various other aspects of the general theory of income and the equations relating to it have been treated by several other writers including Frisch, Hayek, Keynes, Kuznets, Lindahl, Littleton, Myrdal, Pigou, Plehn, Tintner.

to C_1, nor would G_1 necessarily be zero, nor would E necessarily be equal to k.

Thus the essential characteristics of the simplest case of a perpetual annuity are that the M's, C's, G's, and E's are all constants, with values respectively as follows: $M_t = k$, $E_t = k$, $C_t = k/i$, $G_t = 0$.

It is, largely, I believe, the very simplicity of this special case of a perpetual annuity which has introduced confusion and complication into the subject of income in accounting, and therefore into terminology, and therefore finally into tax legislation. In particular, E and M (being in this simple case always equal) are commonly confused.

Even when E and M are numerically equal, they should be distinguished. Two distinctions may here be emphasized:

(1) E accrues continuously—from the moment after any M occurs up to the moment before the next M occurs; whereas the M's occur discontinuously.[13]

(2) The continuous accrual of E means a continuous *increase* of C between M's; whereas the payment of the next M causes C to *decrease* suddenly by that amount. The value of the annuity or *rente* is said to go then "ex-interest." Thus a graph of C is a saw-tooth curve, as in the preceding chart, first *rising gradually* and then *falling suddenly*. This is true, of course, not simply of perpetual annuities but in general, E always causing a *gradual rise* and M always causing a *sudden fall* of C.

In the simple case here supposed, the *rente* may be bought or sold at the same price, k/i, at the end of every year. For instance, it may be bought at this price, kept for ten years, during which $M = k$ is received annually, and then sold at this same price.

We have previously thought of the C's as appraisals, not actual payments. But if, as just supposed, a *rente* is actually bought and sold, the sums paid are also M's—as truly as are any other payments. Thus, the purchase and sale just supposed amount to carving out of the perpetual annuity a ten year segment, comprising twelve M's altogether, including the initial purchase and the final sale.

The supposed series of twelve M's, in connection with a perpetual annuity, is closely imitated in the case of a ten-year bond bought at par. In America, perpetual *rentes* are not in use as they are in France. But bonds, which amount to nearly the same thing, are very widely used. The purchase of an n-year bond may be thought of as substantially the purchase of a perpetual annuity coupled with its resale at the end of n years.

Equation (1), applied to the bond, gives

[13] Theoretically, of course, the M's *could* conceivably occur continuously or nearly so, and services in kind sometimes do flow continuously (e.g., electrical, gas, water, services); but E *must* accrue continuously.

$$C_0 = \sum_{t=1}^{t=n} kv^t + Pv^n,$$

where P is the final payment, called "*p*rincipal," in addition to the final k. This is general;[14] but, in order to obtain the greatest simplicity, we wish here to assume a special case, namely, the case where the bond is bought at par ($C_0=P$). Therefore,

$$P = \sum_{t=1}^{t=n} kv^t + Pv^n.$$

Since we are assuming that i remains forever the same, we may suppose that, at the end of the n years, the bond will be refunded with an exactly similar bond. Or, if not *refunded* in the sense of being simply exchanged for, and replaced by, the new bond, it may be extended in almost exactly the same way if the "principal" P when paid is immediately *re-invested* in an exactly similar bond. In either case, the first bond is replaced at maturity by another for which

$$P = \sum_{t=n+1}^{t=2n} kv^t + Pv^n.$$

By substituting this value of P in the last term of the former equation, it evidently becomes

$$P = \sum_{t=1}^{t=2n} kv^t + Pv^{2n},$$

representing a bond running twice as long as the original. Repeating this reasoning, we may still further extend the term of the bond successively as

$$P = \sum_{t=1}^{t=xn} kv^t + Pv^{xn}.$$

As the factor x increases, the remainder term Pv^{xn} decreases, since v is a proper fraction, and approaches zero as a limit. So that ultimately

$$P = \sum_{t=1}^{t=\infty} kv^t,$$

which is the formula for a perpetual annuity.

[14] Another form of this general equation, and one more convenient for computing the value of a bond, is: $C_0=k/i+(P-k/i)v^n$.

We have seen that, in this case, $i = k/C_0 = k/P$. There are other ways[15] of proving the same results.

As to G, it is zero for each yearly saw tooth, except the last, when it is negative by an amount equal to the "principal." As to E, it is equal to k for every year (including the last).

We have just seen that a bond can be converted into a perpetual annuity by perpetual refunding or reinvestment, as the "principal" is made to recede into the infinite future.

If we move the principal in the opposite direction—from future to present—we obtain a third special case, namely one where there is only one M, and that one M payable immediately. This is the simplest case of all and needs no discussion.

A fourth special case is obtained by moving into the future not only the principal, P, but every other item (namely each "interest" payment, k) all those deferred items accumulating at compound interest. Deferred to infinity, the resulting aggregate becomes infinite. But its present value still remains C_0. This is the case of a savings-bank deposit accumulating indefinitely without either new deposits or any withdrawals.

In this case there is no M until the end of the postponement, when it may be received as one lump sum. The C increases yearly by $G = E = Ci$. This yearly increase, Ci, itself increases as C increases.

The case of a savings-bank account is no exception to the proposition that C_0 is the discounted value of future M's, as in the case of a promissory note. In fact, a savings-bank account is fundamentally a part ownership in a multitude of promissory notes, the mortgages held by the bank.

[15] One is to proceed backward from the terminal date of the bond. At that time there is a total payment of $k+P$ so that, one year before, the value of the bond, C_{n-1}, is $(k+P)/(1+i)$. What this value is depends on whether k/P is equal to, greater than, or less than, i. If $k/P = i$, evidently the above $(k+P)/(1+i) = P$. That is, $C_{n-1} = P$. By similar reasoning it may be seen that $C_{n-2} = P$ and so on backward, including $C_0 = P$. Thus, the condition $k/P = i$ is *sufficient* to make $C_0 = P$.

Moreover this condition is also *necessary*. That is k/P cannot be $>i$ nor $<i$ if $C_0 = P$. For, if it could be true that $k/P > i$, similar reasoning to that above would show that all C's, including C_0, would exceed P, which is contrary to our assumption that $C_0 = P$. The opposite would hold if $k/P < i$. It follows that, if $C_0 = P$, then k/P must equal i.

This method of proof, unlike that used in the text, does not need to assume that i remains unchanged after the termination of the bond.

The reader will recognize k/P as the "nominal" rate of interest of the bond, while i is the "rate realized" by the investor. What has been proved is that, if these two rates are equal, the value, C_t, of the bond will always equal P, but not otherwise.

CHAPTER IV: NORMS FOR MONEY FROM INVESTMENT

We have mentioned four special cases. Of these the perpetual annuity is ordinarily considered the "Normal" case. It often exists merely as a theoretical ideal. As has been seen, a bond may, with certain provisos, be thought of as a sort of segment of a perpetual annuity. The last payment of a bond, M_n, is, as we have seen, composed of two parts: the usual annual payment k, and the remainder, M_n-k, a larger sum called "principal."

This principal is commonly thought of as something unlike the annual k's and not to be "used as income" but to be reinvested. To fail to reinvest it would, it is thought, constitute a gross impairment of capital, something highly abnormal.

Where there is such a punctilious regard for maintaining capital unimpaired, it often leads to legislation and customs requiring trustees and others to reinvest even some of the "interest" payments, k, in case $i<k/P$, which implies that $C_0>P$, that is that the bond sells "at a premium," or above par. In this case, not the actual M's received (the k's), but only the earnings, C_0i, should, strictly speaking, be considered as the normal perpetual annuity. The difference, $k-C_0i$, which is a fixed percentage of C_0, must be reinvested every year. This difference is "savings" and is measured *relatively to C_0i or E, as ideally constant earnings.*

On the other hand, if $C_0<P$, the bondholder is entitled, according to the same standards, to use "as income," not only k, the actual receipts, but also C_0i-k by, say, selling some of his bonds and "spending" the proceeds, "as income."

If every bond bought at a premium could be accompanied by its own special depreciation fund into which the sum $k-C_0i$ were paid annually the actuality would agree with the norm; although, even then, it would not be *the bond alone but the combined bond and depreciation fund* which yielded the perpetual annuity of C_0i per year. This *combined* property would then have a constant capitalized value of C_0; and this value would be made up of two parts, the value of the bond (diminishing every year) and the value of the depreciation fund (increasing every year by exactly the same amount). In this way the perpetual annuity is used by accountants as a "yardstick" for comparison with, or even for correcting, the actual original M's.

But if, as is usually the case, no such separate depreciation fund exists, no such perpetual annuity definitely emerges unless some other form of smoothing out the irregularities of the M's is employed. All that such bookkeeping "funds" amount to is to change the classification of the asset and liability items in such a way as to carve out the required constants. Much of the ingenuity of accountants is spent on

such smoothing devices—depreciation funds, sinking funds, reserves for obsolescence, and so on.

The same idea of using the perpetual annuity as a norm is applicable not only to bonds but to any other series of income received, M_t. Starting with C_0, and the ideal perpetual annuity, C_0i, the deviation of every M from C_0i is supposed to be reckoned. The amount of this deviation is to be reinvested when M_t is too big.

No objection to this procedure need be made. But merely to *reckon* such corrections does not make them effective. Only so far as they are made effective does the smoothing actually occur. And in practice it is never a perfect smoothing. In fact, scarcely anyone would care to keep to such an inelastic series of M's perpetually.

In short, M is what is actually received; E is what "ought" to be received, the norm.

It is interesting to note that the stigma of depleting capital has not usually been applied in reverse. On the contrary, to increase capital by "saving" has usually been regarded with favor. Any such "saving" out of the "normal" M_q may be called S_q. That is, savings are here defined by the equation $S_q = G_q$ so that:

$$E_q = M_q + S_q. \tag{15}$$

The reason for introducing the new letter, S, though the old one, G, represents precisely the same magnitude, will be made evident when we discuss income from labor. Earnings and savings will then be redefined to suit a different norm.

It should now be clear that we must make a great distinction between a mere norm and an actuality. It is true that human life is largely repetitive. We eat, ordinarily, every day. We use our dining rooms day by day, our bed rooms night by night, our churches, or movies, or golf links, week by week. We use almost all our consumption goods and spend our money on them in much the same way day after day, year after year. Fundamentally we repeat these uses of goods and the spending of money for them because of two sorts of cycles, that of our bodily metabolism and that of our environment which brings us day and night.

The resulting repetition in our lives is the chief reason why we like to treat large "non-recurring" M's, the P, for instance, of a bond, differently from the recurring items, and to "spread" them over some stretch of time. But none of the repetition or recurrence is ever precise and each person must decide for himself how much elasticity, or departure from so rigid a norm as a perpetual annuity, suits his own particular case.

In practice, and for most natural persons, a series of money receipts,

$M_1, M_2, \cdots M_n$, from a property ought, in general, to be fairly steady but not rigidly uniform. Yet any individual property—be it a stock, bond, real estate, or whatever else—may yield a series of receipts far from steady, provided this unsteadiness is roughly compensated for by other properties. If, for instance, John Smith owns both a mine and a forest, the M's he receives from the mine may, for a time, greatly exceed the norm of constancy until the mine is exhausted while the M's from the growing forest for the same period may be far below the norm of constancy until time for cutting arrives. The two series of M's thus tend to correct each other, and little or no separate correction for each is needed. All our John Smith needs is such an assortment of properties as will bring him in a reasonably steady flow of money from all of them *combined*, and even that total flow will need to vary somewhat according to varying conditions—John Smith's age, the size of his family, illnesses, and so forth.

But most M's from property flow from corporations, and these are set up as "perpetual." Accountants naturally find the perpetual annuity the only natural norm for use in corporation accounting.

One result is that most of our investment securities are patterned on a perpetual annuity. This is true, for instance, of bonds, mortgages, and preferred stock.

CHAPTER V: THE FACTOR OF CHANCE

Up to this point we have assumed that the M's can be, and are, definitely foreknown; but this is seldom true—in fact, strictly speaking, it is never true. There are always chances of receiving more or less than was expected.

To make allowance for these chances, our three hypotheses are now dropped. Instead, we make the hypothesis that the series of future payments M_t are not definitely known in advance but have to be estimated as to amount, as to time of occurrence, and even as to their occurring at all.

In the case of stocks, for instance, the future M's (dividends) are uncertain in every way. To calculate the capital value of future expected dividends, we have to do a good deal of guessing. Therefore, besides the five sorts of original factors, i, M's, C's, G's, E's, we need now to take account of a sixth factor—probability or chance.

There seems to be no best way to do this; the ways used in *The Nature of Capital and Income* were frankly makeshifts. I had hoped, in these 30 years, to observe more progress by others toward a mathematical, yet realistic and usable, treatment of the probability factor. Various efforts have been made in this direction, but the only worth-

while systems still seem to be confined to life insurance calculations, where probability has a statistical foundation.

Fortunately, however, the vagueness introduced by this chance factor does not affect our most fundamental propositions. C may still be called the present value of future *expected* dividends, and E may still be taken as the sum of M and S.

Thus, in a corporation, the undistributed earnings in any year q are S_q, the distributed earnings, M_q, and the total earnings, $E_q = M_q + S_q$. These may be taken as applying either to the total figures of the corporation or to earnings per share of each class of stock.

G_q will be equal to S_q if both are reckoned at the figure entered by the corporation's bookkeeper as addition to surplus. But this book figure will not necessarily agree with market values in the speculative exchanges. Thus G here becomes ambiguous, as there are at least two possible values for it, the increase in market value and the excess of earnings over dividends as computed in corporate accounting.

Let us now see what the chance element does to some of the equations which we have seen to hold true under our original, chanceless, hypotheses.

The chief change is in equation (1). One summary way of making this new adjustment is simply to insert, on the right-hand side, a multiplier, X, to take account of chance, letting v have the same meaning as before and letting M_t be the "expected," or "most probable," receipt at time t. In other words X may be defined as the ratio of calculated capital value to the actual capital value, the former being calculated from any arbitrarily chosen series M_t, designated as "expected." This gives us

$$C_q = X \sum_{t=q+1}^{t=n} M_t v^{t-q}. \tag{16}$$

Equation (16) reduces to equation (1) when uncertainty is removed and $X = 1$.

This formula throws the whole effect of the chance element on to this one factor, X. But we cannot specify fully what sort of a function X is. Any discussion of this must be purely academic. The value of C will be found in the market by rule of "thumb" or "by cutting the Gordian knot" regardless of any formulae. We can scarcely expect to find any formula for determining the value of a stock, on the basis of expectations, which will be comparable in usefulness to the tables employed by brokers for bond values in relation to the rate of interest, the nominal interest, and the time to maturity.

If the probability of big future dividends is great, C_q is high relatively to the last reported earnings. If, on the other hand, there is great risk of low, or zero, dividends, C_q is low relatively to those earnings.

As to C_q, market value is all important. Book value is of little or no significance for any purposes here considered.

But this is not the place to pursue so intricate and baffling a subject in detail. We need only emphasize the fact that the introduction of the element of chance does not destroy, but rather adds to, the conclusion that C is not, in general, the discounted value of the series of future expected E's but only of the series of future expected M's.

Similarly, equation (4) needs modification. We can scarcely say that $E_1 = C_0 i$. But we can come very close to this equation if we replace i by j, the earnings-price ratio. That is, while i relates only to theoretical certainty, its substitute, j, takes account of the uncertainty pertaining to each individual stock as recorded in its market price. This "j" may be greater or less than the "i"'s used in bond calculations.

Instead, therefore, of equation (4) we have (for the first year)

$$E_1 = C_0 j_1, \tag{17}$$

and the analogous equation for any other year.

To summarize, as to equations (1)–(15), we see that:

(2), (5), (15) remain true;
(1) is no longer true, but is superseded by (16);
(3) remains true provided G be taken as equal to S;
(4) is no longer true but is superseded by (17).

None of the other 12 equations is true under our present hypotheses. But they can all be replaced by somewhat analogous equations just as (1) and (4) were so replaced.

Where chance enters, the need of smoothing devices is greatly intensified. Besides the methods of smoothing already noted, there are insurance and hedging. Insurance substitutes a steady annual "premium" payment in place of a (possible) sudden large outgo—the loss insured against.

The common norm for stocks, both market values and book values, is a constant C, or at any rate an unimpaired C. Starting with C_0, the capital may rise above C_0 but it is not supposed to fall below it, although after rising above (that is, after the accumulation of a surplus) dividends may be declared out of this surplus, even if no addition to surplus is currently earned, provided no more than the surplus is used up—in other words, provided the original capital C_0 is unimpaired. In general, the corporation norm is this maintenance of unimpaired capital. Any earnings above this are "available for dividends," which means permissible under the norm. These ideas are further evidence that earnings always relate to possible, and not necessarily actual, M's.

This norm of an unimpaired C_0 for stocks under the hypothesis of uncertainty corresponds roughly to the usual norm already discussed

for bonds under the hypothesis of certainty. But for stocks there is nothing to correspond closely to a perpetual annuity as a norm.

Of course, the hypothesis of certainty is never absolutely realized. Chance is always playing some part. C, in particular, being always the discounted value of uncertain future receipts, is subject to chance fluctuations. Nor is the hypothesis of an eternally constant i ever fully realized.

The various influences tending to affect C may now be summarized as follows:

(1) C tends to increase with the approach in time of future expected M's;

(2) C tends to decrease with the payment of each M;

(3) C tends to increase with a decrease of i (or j) (and reversely);

(4) C tends to increase with each new expectation of increased future M's (and reversely).

In all cases, C is supposed to be measured in terms of money. If the purchasing power of money varies, C will, in terms of purchasing power, be affected according as the variations in that purchasing power affect the several magnitudes above mentioned. The effects, in extreme cases, become very complicated. They also play havoc with norms.

CHAPTER VI: FORMULAE ON MONEY FROM LABOR

We come now to the M's from labor—wages, salaries, professional fees, and commissions, received in return for work.

In this case it is unusual to derive from the M's the corresponding C's, G's, and E's. Theoretically this can be done, of course. In fact, C has sometimes actually been calculated, especially by life insurance statisticians like Louis Dublin of the Metropolitan Life Insurance Company. Two generations ago this was also done by William Farr, the English authority on vital statistics. J. Shield Nicholson, the Scotch economist, also attempted it. In a small way I did. The net result generally turns out to be that the capital value of money received from the labor of the existing population—or the value of the "vital capital"—is about five times the value of all other capital put together—land, land-improvements, machinery, equipment, and other inventory.

But such calculations of the capital value of human working power are generally regarded as curiosities and have little practical application except possibly as a check sometimes on the over-insurance of a person's life.

The capitalized value of the future salary, or other money received from work, usually shrinks as time goes on. The value for a man when fifty years old is usually less than it was when he was forty and had

more years of work to look forward to. In that case, G, the "capital gain" is negative. This fact of negative gain is usually neglected; but such neglect is inconsistent with the usage of accountants in trying to keep unimpaired the capital of corporations. If the accounting of salaries and wages were on all fours with the accounting of a bond, or of other items of property or corporate accounting, only a small part of the money received for the work of an elderly man could be considered "income." To spend it all on living expenses would be to "eat up his capital."

This inconsistency of usage as between money received from investment and money received from labor is, I suspect, fundamentally due to the fact that accountants' norms began with "artificial persons" such as a corporation, whose life is "perpetual," while an individual's life is not. At any rate the fact is that we now have two conflicting traditions regarding the definition of earnings, one for individual bookkeeping and the other for corporation accounting. The bookkeeping tradition for the individual is the older of the two and presumably originated in the idea that, if a man's wages support him through his lifetime, or even only through his working years, it is perfectly legitimate for him to use them up. The next generation can take care of itself—or even take care of him when superannuated. He regards the money which he receives and which we shall, as before, call M, as his "earnings" which we shall also call E'. Here the two concepts, income and earnings, agree.

On the other hand, the newer tradition, that of corporation accounting, has proceeded on the theory that the corporation is to survive the present generation of stockholders and to serve their successors perhaps perpetually. Thus it comes about that the new norm adopted by property owners is a perpetual annuity, whereas the old norm, so far as any existed, was, vaguely, a life annuity.

There would be no inconsistency between the two norms if the salary earner were assured of living forever and always earning the same salary, M. In that case we would again be dealing with a perpetual annuity; and the capital gain, G, would be zero as before. That is, earnings would then equal salary or wages even when earnings are calculated as in equation (3) for property.

In this hypothetical case, therefore—the case of perpetual human life—the two terminologies (for property and labor) would agree. But in the cases actually existing, they disagree, so far as earnings are concerned. The disagreement is because of a conflict in norms. Any "norm" is an ethical concept of what ought to be, not of what is.

Out of this conflict of norms come great confusions. Since labor's E' and M are identical, it is difficult for many not to identify also the E and M of property.

While depreciation of the capital value of human working power has been ignored by accountants, or rather regarded as outside of the scope of their professional work, nevertheless the need of providing for one's old age and funeral expenses and for giving the children, if any, a good start in life has led to a general precept to "save" out of the E''s and put the savings by for a "rainy day." If a year later the "rainy day" for which these savings were put aside arrives, this sum, with interest, is taken out of the savings bank. The entire sum may then properly be "used as income" without fear of reproach for eating up capital.

If a wage earner or salary receiver saves out of his E''s a certain weekly sum, S', investing this in a savings-bank account or otherwise, he reduces his net receipts to $E'-S'$. But there is no set standard for S'. In fact, a common norm seems to be $S'=0$—if not below zero in this age where thrift is no longer generally glorified as a virtue.

Thus the two norms—one of earnings from investments, the other of earnings from labor—collide in our income-tax laws and policy.

One obvious reason for such collision is that the norms for the M's from labor are so vague that, until the advent, in the last century, of life insurance, there was no good statistical measure of the capital value of one's working power or of life annuities. It would not be surprising if, within another century, new and more definite norms should develop as a result of the actuarial calculations of life insurance companies, old age insurance, and of the growing vogue of life annuities.

It is evident that all norms are somewhat arbitrary. They differ and, in many cases, ought to differ. The norms for a corporation should differ from those of an individual and those for different individuals should differ from one another.

Any type of norm is theoretically possible to suit idiosyncrasies or special circumstances; and after the norm has been set up, savings may be defined with respect to that norm.

One important factor affecting the norm of a wage earner, or salary earner is the number of children or other dependents to be supported. Supporting children out of money from labor corresponds roughly to reinvestment out of money from investment. The latter aims to keep up the capital value of property; the former aims, roughly, to keep up, from one generation to the next, the capital value of the family's labor power.

We have seen that the difference between the usual property norm and the usual labor norm results in slightly different concepts in the two cases. Those for property, as has been noted, are i, M's, C's, E's, S's, G's; and those for labor are:

(1) Receipts of money (the M's) paid for work. These M's have exactly the same significance as the M's from property;

(2) Earnings, E', identical with the M's;

(3) Savings, S', out of E'. These savings (S') are somewhat analogous to the G and S of property. In fact, a definition may be formulated which will include both sorts of savings. If we assume the money-from-labor norm to be $E'=M$ while the money-from-investment norm is $E=M+G=M+S$, then, just as the wage earner saves his S' out of his E', so does the bondholder save his S out of his E, and the corporation save its S out of its E. The only difference is in the E, the norm adopted.

We may perfect the analogy between the equations for labor and those for investments as follows: If we regard the laborer's savings, S', as a *negative* receipt or outgo, we may consider $M-S'$ as his net receipts, M'. Thus

$$M' = M - S'. \tag{18}$$

If we assume, as we have, his norm E' to be identical with M, i.e.,

$$M = E', \tag{19}$$

then we may substitute E' for M, and obtain $M'=E'-S'$, or,

$$E' = M' + S', \tag{20}$$

as compared with $E=M+S$, for investments. In this equation the M' from labor is somewhat analogous to the M from investments, and the S' for labor is somewhat analogous to the S for investments.

But S and S' and also M and M' have certain differences.

In the case of a corporation only the dividends, M, are actually paid, while the rest of the earnings, the savings, S, or undivided profits, are simply withheld from payment; but in the case of the wages earned, the whole earnings are paid to the workman and then the savings are paid out again into the savings bank.

The analogy could conceivably be made more complete in either of two ways. The corporation might first declare all its earnings in dividends and then the recipients reinvest some of them as savings, just as now the wage earner first gets all his earnings in cash and then reinvests some of them as savings.

Or, reversely, the wage earner's savings might, by prearrangement, be withheld and put in a savings fund by his employer without first being paid to him, just as now the corporation withholds undivided profits.

To make any closer analogy we should have to change either or both norms. Thus, we could conceivably set up the same norm for labor as we did for property; then proceed to compute the capital value of a man's future salary or wages and aim to keep this value constant by reinvesting. Or, contrariwise, we could alter the norm for a corporation

to make it resemble that for labor. A mining company would then not need to pay anything into a depreciation fund or a reserve for mine depletion, but could declare dividends M as fast as the mine exploitation permitted. "Savings" would then be considered to be any "savings" out of such M's, even if these were far from sufficient to maintain unimpaired the company's capital. Savings would not then be identical with capital gain for investments any more than for labor.

Such a norm for a corporation's earnings would not be so startling in the case of a mine as in the case of an industry. For a mine is like a man in that it hasn't perpetual "life." To force it into the accountants' norm of a perpetual life, the depletion reserve has to be invested in something else than the mine itself.

But, while it would thus be quite possible to harmonize corporate earnings and labor earnings, by adopting either of the two norms uniformly for both, there is really no need of doing this. On the contrary, the difference between the traditional labor-norm and the traditional corporation-norm, as expressed in the different meanings for "earnings" in the two cases, is based fundamentally, as already hinted, on the difference in the length of life of an individual and that of a corporation. In both cases the norm may be said to be a "life" annuity.

We have seen, therefore, that the two conflicting usages as to the term "earnings" are a matter of practical norms and ethics and not a matter of economic theory. This is evidenced by the ethical connotation of the term "earned" in other connections. "Earned" means "deserved" or "merited." The phrase "unearned increment," for instance, was a phrase calculated to discredit that increment and to prepare the way for taking it away from those who did not deserve it.

By this connotation the income earned by property is that which the owner can rightfully or properly "use as income." If the owner takes more, it is unearned, undeserved.

Whatever norm is used under the title "earnings," the M and S for property, or the M' and S' for labor are the two constituent parts of those earnings and are mutually exclusive parts. If either is increased the other must be decreased by exactly the same amount. You cannot have your cake and eat it too. Savings when taken out of income are no longer income but capital.

CHAPTER VII: SUMMATION OF SERVICES

We have now discussed, in particular, money receipts from property and money receipts from work, and have noted wherein these two differ.

The sum, or aggregation, of the services received by John Smith includes not only services in *m*oney ($\sum M$) but services in kind—"*r*eal"—

($\sum R$), translated (by appropriate price factors) into terms of money, and, finally, psychical services or satisfactions ($\sum Y$), also translated into terms of money.

The services in kind especially considered in this chapter are those rendered by "consumption goods"; such services are the shelter of John Smith's dwelling, the use of his furniture, the wear of his clothes, the use of his food. Psychical services are rendered by his "body-mind"; such services are the enjoyments following the use of his food, clothes, furniture, dwelling, and other services in kind.

It will be seen that the services in money, the consumption services in kind, and the psychical services come in succession and articulate with each other. Yet there will be no double counting by including them all in one sum, provided we are careful to include those which are negative as well as those which are positive. We simply take the algebraic sum of all the services rendered by all the assets of John Smith. This sum, during a specific period of time, we call John Smith's income and denote by "I."

In such a complete accounting there are many negative terms. For instance, every M service which is credited to some asset of John Smith, such as a bond, must also be debited to that central asset called "cash" and, reversely, every M service credited to cash must be debited to some other asset such as John Smith's house, which, let us suppose, is repaired at the cost to him of money.

In the same way, every R service, such as shelter, must be credited to some asset, such as the house, and be debited to the body-mind of John Smith sheltered thereby.

Thus every M is both positive and negative according to the asset to which we relate it—say the bond which yields it and the "cash" which takes it. It is an "interaction" between the bond and the cash. Similarly, every R is both positive and negative in an interaction.

Finally, in a complete picture, there are negative Y's, such as the irksomeness of labor.

Hitherto, for brevity, the M's have been algebraically added together irrespective of whether their numerical values were positive or negative. But henceforth the negative terms will be separately designated. This procedure will be adopted not only as to the M's but also as to the R's and the Y's. The separate designations for the positive and negative M's from investments will be made by the subscript letters r and d; thus $M_r - M_d$ means the money *r*eceived from investments less the money *d*isbursed into investments.

Evidently $\sum M_r - \sum M_d$ is what has hitherto been called $\sum M$, i.e., the algebraic sum of money receipts (money disbursements being negative receipts), or, $\sum M = \sum M_r + \sum(-M_d)$.

As to the R's, since we shall need to treat these only as a single group and enter them twice—once as a positive sum and once as an equal negative sum—we shall only distinguish the two by prefixing a "+" and a "−."

As to the Y's, the *p*ositive items will be distinguished as Y_p and the *n*egative by Y_n.

These different ways of distinguishing the positive and negative items (and the way to be used later with respect to "cash") will be easier to remember and less confusing than a rigidly uniform system of notation.

Let us now classify the component parts of John Smith's income, as above defined, according to sources:

[1] $\sum M_r - \sum M_d$ comprises the sum of all *m*onies *r*eceived within the year from investments or property, like stocks, bonds, real estate, if rented for money, etc. (including any proceeds of sales within the year) less the sum of all *m*onies *d*isbursed (including the purchase price of said property, or such part thereof as is paid within the given year). That is [1] is the net money received from investments.

Included also in class [1], are the monies from (or to) *negative* assets. A debt contributes an M_r when it furnishes the money borrowed by Smith and requires an M_d when repaid (interest or principal). M_r includes all money receipts whatever, (even gifts and bequests) not included under M_w below, and M_d includes all money disbursements not included under M_s below.

[2] $\sum M_w - \sum Y_n$ comprises the sum of all *m*onies received for *w*ork —Smith's salary, wages, fees, or commissions—less the sum of all its psychical cost—its "irksomeness," its "headaches," its "backaches," or, more simply, "labor," all translated into terms of money.

[3] $\sum R - \sum M_s$ comprises the money value of the "*r*eal" services of consumption goods, less the *m*oney *s*pent on them. That is, [3] is the net value received from consumption goods. The money costs, M_s, include any costs of purchase occurring within the year, such as any money paid for a new automobile or a new residence.

The special characteristic of class [3], "consumption" items, is that these services, though objective, immediately precede the subjective or psychical services for which alone all M's and R's exist.

[4] $\sum Y_p - \sum R$ comprises the psychical satisfactions translated into terms of money enjoyed by John Smith less (the money value of) their cost in the form of the services of the consumption goods mentioned under [3].

The special characteristic of [4] is that it comprises all the final satisfactions for which the antecedent money services and services in kind existed. That is, [4] is a net value received from the "body-mind."

[5] $\sum M_{cr} - \sum M_{db}$, (which—for short—will also be called *m*) comprises the *m*oney *cr*edited to the cash, (that is, the money paid out from Smith's cash assets—from his pocketbook, till, cashbox, or checking account) less the *m*oney *deb*ited to cash (that it, the money paid into said repositories of cash). That is, [5] is the net value received from all "cash" assets.

Smith's cash renders a service, as accountants correctly maintain, every time it is used to furnish him cash—to yield up cash—and renders him a disservice every time it takes in cash. A store of cash is like a gold mine, credited with what comes out of it and debited with what goes into it.

[6] *U* (*u*nclassified) comprises all items not included in the preceding five classifications. These will be ignored, or their sum assumed to be zero, until we have fully discussed the first five categories.

We have now before us the broadest possible picture of Smith's income, defined as the algebraic sum of all services received by him, positive or negative—for this summation is indiscriminate, regardless of norms, includes "capital" expenditures along with current expenditures, and is uncorrected for depreciation or anything else. The only corrections will be those automatically made by negative terms offsetting positive terms.

Note that $\sum R$ under [3] is identical with $\sum R$ under [4], the two representing opposite facets of the same "interactions."

Smith's income (complete sum of services, *I*) is therefore:

$$
\begin{aligned}
(21) \qquad I = {} & (\textstyle\sum \dot{M}_r - \sum M_d) && [1]\\
& + (\textstyle\sum M_w - \sum Y_n) && [2]\\
& + (\textstyle\sum R - \sum M_s) && [3]\\
& + (\textstyle\sum Y_p - \sum R) && [4]\\
& + (\textstyle\sum M_{cr} - \sum M_{db}) && [5]\\
& + U. && [6]
\end{aligned}
$$

Summarized in words, equation (21) is

$I =$ [1] the money received from all John Smith's properties less the money expended on them;

+[2] the money received for John Smith's *work*, less its psychic cost (measured in money);

+[3] the services rendered by his "consumption goods" (measured in money) less the money spent on them;

+[4] certain services rendered by Smith's own body-mind to his stream of consciousness, less their "real" cost (both measured in money);

+ [5] the services rendered by "cash," namely its use in furnishing money, less its disservices in absorbing money;

+ [6] unclassified (measured in money).

In equation (21) the terms are grouped in a manner appropriate for our modern (money) economy and for the particular purposes of this study. In this modern world of ours all the terms except [1] and [2] are usually very small. Yet the equation holds good even when the use of money is relatively little—or even wholly lacking.

Of course, anyone who so wishes is free to rearrange the order of the terms. Thus, if it is desired to include under [1] and [2] any items in kind, this can be accomplished by transferring them from R in [3]. Thus, the use of a parsonage by a country clergyman as part of his salary may, if desired, be segregated from R in [3] and transferred to [2]; i.e., added to M_w. The same may be done with the "board and lodging" given as part pay to a domestic servant. But, except for a few such cases, all the recompense for work—in this country, at least—is paid in money, M_w.

Cases in which returns from property are partly in kind are even rarer than the cases just noted in which workers are so paid. The only notable case is that of the farmer. His property still yields him some of his food, though it no longer, to any appreciable extent, yields him his clothing, or his other R's. If desired, a variant of equation (21) might be formulated for farmers by transferring from [3] to [1] such food items among the R's as are consumed on the farms.

Originally, of course, instead of receiving money to be spent, everybody received goods in kind. In this respect everybody was a sort of farmer. But even where this system still survives in part, equation (21) holds true. And it will obviously make no difference to the net income of John Smith whether all the items in kind are kept in [3], as in equation (21), or some are transferred to [1] and [2]. Such transfers will usually result in diminishing the difference between $\sum R$ and $\sum M_s$, i.e., diminishing term [3], and increasing terms [1] and [2].

In the multitudinous sum expressed in equation (21), most of the items form twins, or couples; each couple being zero, consisting of two equal and opposite items. Such a couple constitutes an "interaction" between two assets (or two liabilities or an asset and a liability) of John Smith. Every M or money item may be so coupled with a mate of opposite kind. So may every R item. If each couple, being zero, is dropped out, only psychical items will survive, the "psychic" income of John Smith. Every money transaction is an interaction between "cash" and something else. There can be no uncancelled *money* item in the entire aggregation I. Even a gift of money to Smith must be credited to the source of the gift and debited to cash. An analogous double entry must be made in the case of a gift of money *by* Smith.

It follows that $\sum M_{cr}$, the sum of all the money used from "cash" as a source, must be exactly equal to the sum of all the expenditures in cash debited to other categories, i.e., to $\sum M_d + \sum M_s$; also that $\sum M_{db}$, the sum of all the money absorbed by "cash," must be exactly equal to the sum of all the money received from other categories, i.e., to $\sum M_r + \sum M_w$. So we have:

$$\sum M_r + \sum M_w - \sum M_{db} = 0, \tag{22}$$

and

$$\sum M_d + \sum M_s - \sum M_{cr} = 0. \tag{23}$$

Services in kind are also all double-entried and may therefore all be cancelled in couples. Like the items in money, they are, in every case, interactions between two assets (including negative assets) to be credited to one and debited to the other. That is, as above indicated, the R's of [3] are, identically and individually, the same as the R's of [4], the one set being preceded by "+" and the other by "−."

If we thus cancel out all the R's and also cancel out all the M's, as per equations (22) and (23), equation (21) becomes

$$I = \sum Y_p - \sum Y_n. \tag{24}$$

It might be inferred that, since all items but the Y's are cancellable by double entries, they need not have been entered at all. This is true if we are interested only in the net result—"*psychic* income." But equation (21) may also be used to show "*real* income," income *spent*, and "*money* income."

It is clear that, under the grouping adopted, the positive side of a couple sometimes has to be entered in one of the six groups while the negative side is entered in another. It is also clear that to include Y's, the psychical elements, in our bookkeeping, we must include Smith's "body-mind" as an asset, with its debits and credits.

If now we choose to omit Smith's "body-mind" as an asset, with all its debits (including the negative R items) and credits, and then, among the terms remaining, cancel out all couples, there will survive, uncancelled, only the consumption items—like shelter, in other words, the "*real* income" of John Smith (the positive R items).

Again, if we choose to omit the assets yielding consumption or real income, like the dwelling house, and then, among the remaining items, cancel out all couples, there will survive, uncancelled, the "*money* income" of John Smith.

Besides thus deriving psychic, real, and money income by selecting the appropriate group of assets and liabilities, we may select any other

group arbitrarily. Whatever we choose to include in our accounting, the uncancelled items will always be positive, never negative.

In numerically illustrating the items separately specified in equation (21), we must remember that there are twin R's, as indicated by their having identical symbols, and we must also remember that the numbers chosen must satisfy equations (22) and (23). The illustrative figures which follow conform to these requirements:

[1] *property:* $\Sigma M_r = \$30{,}000$, money received from property—stocks, bonds, rented real estate, etc.—less $\Sigma M_d = \$25{,}000$, money expended on property—taxes, repairs, etc.—leaving a net money from property of $+\$5{,}000$;

[2] *work:* $\Sigma M_w = \$8{,}000$ money received from work—salary, wages, fees, commissions—less $\Sigma Y_n = \$2{,}000$, as appraisal of the irksomeness of labor—leaving a net value received from work of $+\$6{,}000$;

[3] *consumption goods:* $\Sigma R = \$15{,}000$, as appraisal of the services rendered by consumption goods—shelter from dwelling, comfort from furniture, use of clothes, food, etc.—less $\Sigma M_s = \$14{,}000$, money spent on consumption goods—leaving a net value received from consumption goods of $+\$1{,}000$;

[4] *Smith's body-mind:* $\Sigma Y_p = \$16{,}000$, as appraisal of its subjective services, less $\Sigma R = \$15{,}000$, as appraisal of its objective disservices, leaving a net from body-mind of $+\$1{,}000$;

[5] *cash:* $\Sigma M_{cr} = \$39{,}000$, money credited to cash, less $\Sigma M_{db} = \$38{,}000$, money debited to cash, leaving a net from "cash" of $+\$1{,}000$;

[6] *unclassified:* U, say \$0.

Accordingly, John Smith's net I, is $+\$5{,}000$ (property) $+\$6{,}000$ (work) $+\$1{,}000$ (consumption goods) $+\$1{,}000$ (body-mind) $+\$1{,}000$ (cash) $+\$0$ (unclassified) $= \$14{,}000$.

Thus we have in one interrelated picture all sorts of income—psychic, real, money, from property, work, consumption goods, body-mind, miscellaneous.

The psychic items, Y's, are of special interest to the theoretical economist, to the philosopher, and to the philanthropist. The latter wants to see the maximum of human satisfaction and cares little for the M's and R's unless they really result in increasing the Y's. If the M's or R's make our John Smith ill, as so often do dark-room tenements or narcotic indulgences, they are a delusion and a snare, however great their market value.

The study of the Y's is a most vital study for economists. Personally I deplore the narrowness of those conceptions of economics which

would leave Y unstudied, which would assume that Y is only the consumer's private business and none of the business of an economist, and which would therefore find no place for a study of whether a person or a nation really obtains, in satisfactions, the money's worth shown on the books.

But this narrowness is disappearing. "Home Economics," which seemed at first an intrusion to many academic economists, is obtaining more attention. Organizations for "consumers' research" are trying to safeguard the consumer against swindling and false advertising. Our pure food laws, sanitary milk laws, sanitary tenement laws, are other examples. Closely associated is the demand for sanitary and pleasant working conditions. All these are engaging the attention of economists.

It is much to be desired that the self-complacency which has so long narrowed the supposedly proper field for economics should be abandoned and the old fences broken down. It is time that more studies be made, better to connect wealth with welfare, which is what "wealth" originally signified. In all fields of knowledge a modern species of specialism is developing which demands the special study of border-line problems. A supremely important border-line problem is how the psychic satisfactions Y_p can be increased, the psychic costs Y_n be decreased, and the entire stream of consciousness of human beings be improved.

If so broad a study must go beyond the boundaries of economics, no harm need be done. But for those who are meticulous about boundary lines, suffice it to exclude from economics proper the aches and pains, the joys and sorrows, and all other elements in John Smith's stream of consciousness which are not attributable to his property or his labor and which therefore have no correspondence to his "real" R's or his "money" M's.

The concepts of money I and real I can merely serve as makeshifts for, or approximations to, the true I, that is the psychical I. The only excuse for them lies in the fact that they are more easily measured in definite units.

It is one of the curious paradoxes of economics that while *theoretically* psychical I is the only perfectly accurate expression for John Smith's complete I as given in equation (21), yet *practically* the psychical factors cannot be independently measured at all. It is easy, sitting in an arm chair, to invent the illustrative figures, \$16,000–\$2,000, for $\sum Y_p - \sum Y_n$, but no statistician, accountant, tax-legislator, or anyone else would make a serious attempt to measure the Y's for John Smith, much less for the nation.

This paradox lies behind the controversy between the so-called "theoretical" economist and the would-be "practical" student. The

latter complains that economics has wandered into "metaphysics," and demands a return to "cold facts." He would have us forget the "Austrian school" of subjective economics, as incapable of giving us any tangible results. One such writer entitled an article "the final futility of final utility."

The late President Hadley remarked:

> The older political economy expressed its results in pounds, shillings, and pence. They might be true or they might be false, but they were at any rate in a form where they were capable of measurement and verification. . . . The new political economy has substituted a more vague conception of wealth for the more concrete one, and many of its propositions have suffered a corresponding loss of clearness and precision. The mercantile school of economists had measured wealth in terms of money. The first generation of their critics measured it in terms of food; the second and third generations measured it in terms of "commodities"; our own generation measures it in terms of utility. But food is a less definite and tangible measure than money; commodities are a less definite and tangible measure than food, and utility is perhaps the least definite and tangible measure of all. . . . I am disposed to think seriously that the excessive use of psychological terms and conceptions, to the neglect of purely commercial ones, has been the most potent cause to weaken the influence of economists among statesmen and men of the world.

It may be claimed that the present analysis has the virtue of coordinating the money concepts of the market place with other concepts including the subjective concepts of theoretical economics without which the market place would be meaningless.

Before proceeding to discuss the possible approximations available without using any Y terms, it should here be said that we need not wholly despair of some day seeing econometric science conquer, to some extent, the problem of measuring even the Y's. In fact, an attempt has been made by me to take a first step in this direction;[16] and Professor Ragnar Frisch had independently done the same[17] in a different way. He has since made still more notable contributions[18] to this elusive subject.

In fact, most of the data needed for measuring the Y's are supplied by the market place, which is merely the place where people decide what certain proposed satisfactions of their wants are worth in money.

[16] "A Statistical Method for Measuring 'Marginal Utility' and Testing the Justice of a Progressive Income Tax," in *Economic essays contributed in honor of John Bates Clark*, published in behalf of the American Economic Association by the Macmillan Co., 1927.

[17] "Sur un Problème d'Économie Pure," Series *Norsk Matematisk Forenings Skrifter*, Serie 1, Nr. 16, 1926.

[18] *New Methods of Measuring Marginal Utility*, Verlag von J. C. B. Mohr (Paul Siebeck), Tübingen, 1932, 142 pp.

CHAPTER VIII: GROSS PSYCHIC INCOME AND THREE APPROXIMATE MONEY MEASURES OF IT

But, in the present state of our knowledge, it still remains true, as President Hadley said, that psychological concepts are not as helpful or as appealing to practical people as commercial ones.

Let us, therefore, see what our theoretical analysis can contribute toward some practical measure of I.

We must now leave behind us the armchair numerical values which were assigned above to the Y_p's and Y_n's. None of these terms in equation (24) are to be found in any of John Smith's accounts and none are calculable except as inferences from the M and R figures which are to be found in those accounts.

Intermediate between the money figures in Smith's accounts and his psychic income, I, there are consumption goods as supposedly equivalent of both. Let us, then, in our search for some good measure of the true I, namely the Y's, first take the intermediate step of translating the Y's into R's.

Before doing even this, however, we may take a prior step, namely to neglect $\sum Y_n$, the "irksomeness of labor." This is an old phrase but nobody has ever translated it into figures which any accountant would respect. According to one interpretation, we should count John Smith's labor as equivalent to his work, because "at the margin" he is indifferent as to whether one more unit of work is worth the effort, or, in familiar phrase, "worth while." According to this interpretation $\sum M_w - \sum Y_n$ is zero.

Another interpretation is not by such *marginal* valuations of labor and work but by integration from the origin. That is, we assess what theoretical economists have called the "quasi rent," that is, we try to answer the question: What would Smith accept in money as an equivalent of $\sum M_w - \sum Y_n$? That question actually does arise occasionally in practice when a salaried man receiving for his work, say $\sum M_w = \$8,000$ reaches retirement age and is given the option of retiring on, say, \$6,000 or continuing to work. If he prefers the \$6,000 without labor to the \$8,000 with labor, then presumably for him, at that time, his \$8,000 with labor is not worth as much as \$6,000, that is, $\sum M_w - \sum Y_n < \$6,000$, and therefore $\sum Y_n > \$2,000$. If, on the other hand, he rejects the offer of \$6,000, then $\sum Y_n < \$2,000$.

This seems a reasonable method of measuring or approximating Y_n for income-tax purposes. But it is very seldom applicable. Apparently the only serious attention given, in practice, to $\sum Y_n$ is to be found in certain tax laws which tax M_w ("income earned" from work) at a lower rate than they tax the same sum when received from property. But no definite standards have yet been set up for such an allowance.

And, in general, Y_n, though important and worthy of long discussion in theory, is almost completely ignored in practice.

Accordingly, from here on, Y_n will be ignored in our discussion, since we are now looking for a practical measure, if not of $I = \sum Y_p - \sum Y_n$, at any rate of

$$I_g = \sum Y_p. \tag{25}$$

This I_g will be called *g*ross psychic income.

Henceforth our approximations will be directed to obtaining the best practicable measure of $\sum Y_p$, first approximating it in terms of the R's and then in terms of the M's.

The first of these two steps is easily taken by assuming that $\sum Y_p = \sum R$. This is little more than the reasonable assumption that the money value of psychical enjoyment is the same as that of the objective services which bring that enjoyment.

It will be remembered that we agreed to include in Y only the psychical reactions to R's as external stimuli, and not to include any other experiences in the stream of consciousness, desirable or undesirable. Therefore, there is an exact one-to-one correspondence between the R items and the Y items. If, now, we could properly assume that each Y item has the same value in terms of money as the corresponding R item and that the two were simultaneous it would follow that, during a given year, $\sum Y$ would equal $\sum R$. Therefore the only possible sources of discrepancy between $\sum R$ and $\sum Y$ are:

(1) The fact that there is a slight time lag between each R and the resultant Y. For instance, the satisfaction from eating comes after the eating. But in almost every case such lags are very small. The act of seeing a movie (the impinging of light on the retina) and the satisfaction from receiving it are practically simultaneous. Scarcely any important case of a long lag between an objective service rendered by a consumption good and the subjective satisfaction which follows that objective service can be found. And where such a long lag does exist, as, say, in education, the fact that some of the satisfactions from this year's reading will come in later years is nearly balanced by the fact that some of this year's satisfactions came from the similar reading done in previous years. We shall assume that these overlaps will approximately balance.

(2) The fact that John Smith sometimes overestimates or underestimates the satisfactions which he expects from his R's. He may eat a dinner and get no satisfaction, but indigestion instead. He may even have been swindled.

But, in general and "normally," we may reasonably make, as a close approximation, the assumption proposed above, that $\sum Y_p = \sum R$. We

may thus proceed from equation (25) to I_r, a new approximation for $\sum Y_p$, namely

$$I_r = \sum R \tag{26}$$

where I_r will be called John Smith's "real income."

Thus, we have in $\sum R$ a fairly good substitute for $\sum Y_p$. It consists of so-called "consumption," that is, the use of food, the use of clothing (its "wear"), the use of a dwelling (shelter) the services of domestic servants, the personal use of an automobile or other conveyance, other personal transportation, telephone service, the use of the theaters, movies, other amusements, the use of books, and the uses or services of every other consumption good. Numerically, according to the illustrative figures chosen, I_r is worth $16,000.

For any who may question the approximate correctness of the hypothesis that $\sum Y_p = \sum R$, it may be pointed out that, even if $\sum Y_n$ and $\sum R$ are very unequal, we may, none the less, pass from the "psychic" concept to the "real" concept by dropping out the discrepancies between them, just as we dropped out $\sum Y_n$. Whether the discrepancies be legitimately cancelled out or arbitrarily dropped out we have left John Smith's "real income," R. From now on, the "practical" reader will feel on safe ground, freed of any obligation to consider "psychical" income.

(Incidentally it may be remarked that no one includes in "real income" any capital item—any G, or S, or S'. That is, no one claims that savings are *real* income.)

We are now ready to pass still further back—to the M's. We do so by assuming that $\sum R = \sum M_s$. If we could properly assume (1) that the two are simultaneous; (2) that each M_s item exactly measures the R item for which it is spent; and (3) that all R items are bought with M items, it would necessarily follow that $\sum M_s = \sum R$.

These assumptions may be almost exactly realized in the case of a person who lives in a rented furnished house, especially if he also rents a garage and automobile. He then "pays as he goes." In such a case the money he spends on consumption goods will be exactly worth the services of these goods to him save for the following exceptions:

(1) There may still be a lag in time in some cases as, for instance, between the payment for a suit of clothes and its use.

(2) The price paid may later prove to have been more than the goods were worth—or less.

(3) There may be some income in kind without any money passing.

Since $I_r = \sum R$ and since $\sum R$ is, for the present, assumed to be *approximately* equal to $\sum M_s$, we see that Smith's income spent on consumption goods, namely $I_s = \sum M_s$, is an approximation to his true gross

psychic income $\sum Y_p$. In the numerical example, the value of $\sum M_s$ is \$14,000.

While, theoretically, the I_s approximation, given by the last equation, may not be as good as the I_r approximation of equation (26), it is far easier for Smith or his accountants to calculate it; because the constituents of $\sum M_s$ are recorded in some detail in his account book. But $\sum M_s$ consists of a vast multitude of items and most of those spent out of the pocket book, rather than out of the check book, are not recorded completely or with accuracy.

It is possible, however, to reach $\sum M_s$ indirectly by using fewer items. By equation (23), $\sum M_s = \sum M_{cr} - \sum M_d$. Adding equation (22) (after reversing the sides) we have $\sum M_s = (\sum M_r - \sum M_d) + \sum M_w + (\sum M_{cr} - \sum M_{db})$. That is,

$$I_s = \sum M_s = (\sum M_r - \sum M_d) + \sum M_w + m. \tag{27}$$

This represents "income spent" or "spendings."

The last term is m, the change in cash balance in the year. It is ordinarily very small and may often be neglected.

If m be neglected, equation (27) reduces to

$$I_m = (\sum M_r - \sum M_d) + \sum M_w. \tag{28}$$

That is, John Smith's "money" income is the sum of the net money which he receives from his property (which implies, of course, deduction of all money invested during the year) and that which he receives from his work. This money income, in our example, is \$13,000.

Thus we have expressed three approximations to John Smith's *gross psychic I* income as given in equation (25) ($I_g = \sum Y_p = \$16{,}000$).

These are his *real* income, his income *spent*, and his *money* income, as follows:

$$I_r = \sum R = \$15{,}000, \tag{26}$$

$$I_s = \sum M_s = (\sum M_r - \sum M_d) + \sum M_w + m = \$14{,}000, \tag{27}$$

$$I_m = (\sum M_r - \sum M_d) + \sum M_w = \$13{,}000. \tag{28}$$

We may pause here to schematize briefly what might be called the normal case of equation (21), as follows:

Neglect $\sum Y_n$ or assume it to be zero; assume $\sum M_{cr} - \sum M_{db} = m = 0$; assume $U = 0$. Under these special assumptions, it is easy to show that $I = I_g = I_r = I_s = I_m$.

This static picture is what seems usually to be in mind when writers talk of money income measuring real income and real income standing for psychic income. In such a picture the discrepancies have disappeared.

We recur now to the three possible sources of discrepancy between I_s or $\sum M_s$ and I_r or $\sum R$.

As to the second—that price valuations may turn out wrong—it is presumably negligible under ordinary conditions. But the other two sources of discrepancy, the time lags involved, and the possible existence of direct income in kind, cannot so easily be overlooked.

It is true that the time during which money rests in "cash" (pocketbook, till, checking account, or safety-deposit box) is usually too small to be worth considering. Normally it is about a fortnight, though in a depression this may be considerably lengthened by "hoarding." But the time between paying for a consumption good and getting the use of it, for which that payment was made, is, in certain cases, great and subject to great variations. In the case of most foods the lag is, of course, very short. But there are examples in which it is long—for instance, if the food is canned and is put in a storeroom by the consumer. In the case of clothing, the time between buying a suit and getting the wear out of it is usually somewhat longer. In the case of a dwelling, the time between paying rent and getting shelter is negligible, but if, instead of paying rent, John Smith buys the house outright, the shelter he thus pays for is mostly in future years.

It is this time factor which makes the chief practical difficulty in estimating the R of any given year by means of the $\sum M_s$ of that year. Except for this and except for some R items coming to John Smith without the intervention of money, we could assume $\sum R - \sum M_s = 0$ just as we assume $\sum Y_p - \sum R = 0$.

Thus the only important problem still to be solved in order to get a practical approximation to I_r (which we have accepted as an approximation to I_g) is how to calculate $\sum R - \sum M_s$ from the records of John Smith's account books.

CHAPTER IX: A PRACTICAL FORMULA FOR MEASURING INCOME

The question, therefore, is: what, from a practical standpoint, is the best formula which we may expect to get from John Smith's account books for approximating his real income, I_r?

Evidently $\sum R = \sum M_s + (\sum R - \sum M_s)$. The first term, $\sum M_s$, may be found from John Smith's money accounts. As to the second term, although $\sum R$ cannot, by itself, be estimated accurately, yet, curiously enough, the entire parenthesis, $\sum R - \sum M_s$, can, as we shall soon see, be estimated with considerable accuracy.

Let us denote by c, the best estimate of this parenthesis. Then we may write $I_{ba} = \sum M_s + c$. That is, the *b*est *a*pproximation to $I_r = \sum R$ is John Smith's total spendings of money for consumption goods, plus a *c*orrection now to be determined.

But before we show how to estimate this correction-term c, let us substitute the more useful expression for $\sum M_s$ given in equation (27). The result is:

$$(29) \qquad I_{ba} = (\sum M_r - \sum M_d) + \sum M_w + m + c.$$

Of this last form all the items (except c) are usually given with the utmost completeness in any person's account books, namely the money received from investments less the money put into them, plus the money received from labor, plus the depletion of cash during the year. Moreover each of these four items consists of only a few entries.

It remains only to find a formula for c, the best estimate of $\sum R - \sum M_s$. This correction, "c," an estimate of $\sum R - \sum M_s$, may be called the total *use-cost discrepancy*, being the difference between the value of the *use* of consumption goods during the year and the money spent on consumption goods during the year (their *cost*). This total use-cost discrepancy is composed of the individual use-cost discrepancies for all the innumerable consumption goods used by John Smith. But in only a few cases, probably not more than the three specified below, is any particular use-cost discrepancy large enough to be worth estimating. Moreover, the small use-cost discrepancies of several hundred items will largely cancel one another, some being positive and some negative.

Therefore, in estimating "c," we need not take account of all the little use-cost discrepancies. Their sum total may be assumed to be practically zero. Only those use-cost discrepancies whose combined influence on the total is likely to be appreciable need be estimated.

The largest use-cost discrepancy is ordinarily to be found in the financing of shelter. Let R_h be the value of the use of John Smith's *h*ouse during the year and M_h be the money spent on the house—whether rent, total purchase price, or installment payments within the year. Then $c_h = R_h - M_h$ is the use-cost discrepancy for that year, for the house. If Smith pays rent there is, presumably, no use-cost discrepancy. But if he already owned the house, so that he pays nothing for it during the given year, the use-cost discrepancy will be great and positive. If, on the other hand, he pays outright within the year for a new house, the discrepancy is also great but negative. Since M_h is obtainable from Smith's accounts the only question now remaining is how to appraise R_h, the value of the shelter service.

In such cases the appraisal sought must ordinarily be based on the original *c*apital value of the *h*ouse, C_h. There may be many ways of calculating the shelter value per annum from C_h. Whatever they are, the result may be expressed by a multiplier h. In other words, hC_h will be the shelter value per annum. Another factor, h', still to be applied is

the fraction of the year during which the shelter from the house is enjoyed. Thus $R_h = h'hC_h$ and therefore $c_h = R_h - M_h = h'hC_h - M_h$.

To decide on the value of the multiplier h is the particular business of accountants. It may depend on the estimated "life" of the house and, therefore, on the rate of depreciation, the rate of interest, and possibly other factors; and these may vary according to the nature of the house. Probably h will usually be in the neighborhood of 10%. The lot on which the house stands should be included.

So much for John Smith's residence. The second case of appreciable use-cost discrepancy is usually found in *house furnishings*.

In this case a short cut will be superior (especially for tax purposes) to a meticulous application of the preceding process to each table and chair, pot and pan, rug and carpet. Instead of appraising the numerous individual constituents of house furnishing we may lump them together, including jewelry, timepieces, and works of art and call the total capital value of the house *f*urnishings C_f. One summary way to obtain C_f is simply to apply a factor to express the presumed ratio of value of the house furnishings to that of the house itself. But, however C_f is found, to it must now be applied two factors analogous to h and h' above, say f and f', so that $f'fC_f$ will be the estimated value of the use of the house furnishings and $c_f = f'fC_f - M_f$ that of their use-cost discrepancy. If the house is rented furnished no such calculations need be made and the use-cost discrepancy may, as already indicated, be presumed to be zero.

Finally, we may apply the same principles to John Smith's automobile. The cost may be called M_a, *m*oney spent for *a*utomobile, and its use-cost discrepancy may be designated as $c_a = a'aC_a - M_a$. The factor a is available from the tables of automobile dealers. That is, these tables give values at the end of various periods, from which values the corresponding factor a may be computed.

The same principles are applicable to every one of the innumerable goods used by John Smith. But, as already noted, we may assume that all, or almost all, the other discrepancies put together will amount to "next to nothing." In cases of exceptionally wealthy individuals, a private motorboat, or airplane may need to be considered. But ordinarily only the three kinds above, one of them, house furnishings, being a miscellaneous group, seem to require any consideration, and then only because and when they are non-recurring, irregular, and numerically large.

We therefore have:

$$\begin{aligned} c &= c_h + c_f + c_a \\ &= (h'hC_h - M_h) + (f'fC_f - M_f) + (a'aC_a - M_a). \end{aligned}$$

Doubtless the above technique for estimating "c" can be improved upon, if more accuracy[19] is worth attempting. But, in any case, the foregoing method is far simpler, more complete, and more accurate than those now used in estimating income taxes. The method here proposed does not dispense with the need of the accountant's art for spreading non-recurring items over time, but it reduces his spreading work to a minimum and confines it to the R's. No M items need be spread or smoothed. They may be used exactly as they occur and this may be done not only without damaging the accuracy of the final result, I_{ba}, but, on the contrary, safeguarding that accuracy against great possible damage from a maze of individual bookkeeping adjustments. By the method here proposed all these needed corrections are automatically collected and resolved into three corrections, the algebraic sum of which is the one correction c. This is essentially the "total use-cost discrepancy," the domestic analogue of the accountant's allowances for depreciation, obsolescence, depletion, etc. But such spreading over time is vastly simpler when applied to the final domestic accounts than when applied to the numerous anterior business accounts. As Professor Canning well says (in a private letter): "This greater simplicity permits one of the great administrative advantages (to the government) and one of the great convenience advantages (to the taxpayer) of the tax base you propose over the one we now utilize."

By the method here proposed, the money received (M_r) from property—in "[1]" of equation (21)—includes all actual money so received, even "principal" receipts from sale of bonds. Likewise, the money disbursed (M_d) includes "capital expenditures." All the elaborate allowances for depreciation, obsolescence, sinking funds, reserves, etc., however useful for other purposes, are unnecessary so far as estimating John Smith's real income is concerned. For, if he lives up to the requirements which these signify, they will all be taken care of anyway—automatically taken care of by reinvestments, and otherwise. And if, as is *always* the case, he does not entirely live up to these accountants' norms but, for instance, spends for pleasure some of the capital he gets by selling stocks, this not only will be recorded in M_s but ought to be. For the accountant to subtract such items, if Smith does not subtract them himself, is to absolve Smith from sins for which he should be held responsible.

[19] Thus even when, as for food, we assume the use-cost discrepancy to be zero, it would be fair to add another correction, a reasonable interest on the average stock of food, also on the average wardrobe of clothes, and everything else in John Smith's inventory of consumption goods.

CHAPTER X: TWO ZERO ADDITIONS

Up to this point, the residual term U has been ignored. It may still be ignored as equal to zero. For, although it consists of a large number of items, they are all in self-cancelling couples, i.e., "interactions." It can contain nothing else. It can contain no M terms, since all M terms have already been specifically provided for. It can contain no Y terms and no R terms for similar reasons. R specifically includes all the services of consumption goods, i.e., those services in kind just preceding psychical satisfactions.

Yet preceding the R's there are, or may be, more terms than under all the other categories put together. When a farmer threshes wheat or a miller converts wheat into flour or a housewife bakes bread, the operation is at once a service rendered by these people and the other productive agents employed and a disservice rendered by the wheat, flour, or bread costing said service. These innumerable but self-cancelling terms are to be listed under U.[20]

We may put U equal to $\sum A - \sum A$. This must be zero; for it consists of pairs, or couples, of equal and opposite terms, each representing an *a*nticipatory service or disservice—an interaction anticipating the R's to come.

The A items are, like the R items, "in kind" but the two differ. The R's are the last links in the chain of objective services while the A's are earlier links in the chain.

John Smith has been pictured as a typical city man whose income comes mostly from M's. But, as was stated at the outset, the formulae are general and are equally applicable to farmers or others whose incomes are largely in kind.

Take the most extreme case, where a farmer is self-sufficient, receiving no money whatever and spending no money.

Then, in equation (21), all the M, or money items become zero so that we have:

$$0 - 0 \qquad [1]$$

$$+ 0 - \sum Y_n \qquad [2]$$

$$+ \sum R - 0 \qquad [3]$$

$$+ \sum Y_p - \sum R \qquad [4]$$

[20] The U terms have importance in the study of the productive processes including all the processes prior to consumption. Therefore, the U terms have more importance for corporations and other artificial persons than for natural persons. For many of the productive processes find a place in corporation records even if not translated into any money equivalents. And they are often so translated when a corporation has several departments, one providing goods to another, and sets up a separate accounting for each department independently.

$$+ 0 - 0 \tag{5}$$

$$+ \sum A - \sum A. \tag{6}$$

This reduces, of course, to $\sum Y_p - \sum Y_n$ just as before. If we neglect $\sum Y_n$, as before, it reduces to $\sum Y_p$; and if $\sum Y_p - \sum R = 0$, it becomes, as before, identical with $\sum R$.

So far there is nothing peculiar in the present case. But we cannot, in this case, put [3], i.e., $\sum R - \sum M_s = 0$. Therefore equations (27) and (28), based on that assumption, cannot, in such a case, serve as a close approximation to the true I. The use-cost discrepancy is too great. In fact the use-cost discrepancy becomes the whole thing. To obtain a close approximation, there would have to be appraisals not only of dwellings, automobiles, and furniture, but of all enjoyable services. Only by such appraisals can we reach such a farmer's true income.

But probably no such farmers—farmers who use no money at all—remain in the entire world, certainly not in the civilized world, and most certainly not in the United States. In fact there are few who do not use money for obtaining a very large part of $\sum R$, their "real" I.

So much for the income of John Smith, a "natural" person, including the $U = 0$. There remains the problem of the "I" of a corporation or other "artificial" person.

Evidently, as a double entry shows, the corporation as such has no income. It is only when it is thought of as consisting of its stockholders that it can be regarded as having income.

Since the personal accounts of these same stockholders show their income from the corporation, we cannot count the corporation's income separately. That would be double counting. The corporation accounts can be used by the government to check up on the stockholders' reports and the corporation can be asked to pay taxes *for* the stockholders and others on the practical principle of "stoppage at the source." But the corporation, of itself and apart from its stockholders, always and necessarily has an income of exactly zero.

The corporation is a bookkeeping dummy. If it deals also with other artificial persons, these merely carry forward the process. But there must be real persons somewhere in order to receive any net income; for this income, as we have seen, does not, in the last analysis, consist of money payments but of satisfactions. Corporations as such cannot feel satisfactions; for "a corporation has no soul" to feel with. But any corporation or artificial person may accumulate capital gain.

While no artificial person can have psychic or real income, every natural person has a total net income every year for the simple reason that he cannot live a year without it even if his only income is a dole or allowance from the poorhouse.

CHAPTER XI: INCOME TAXATION

There are many applications of the foregoing theory of income—to economic theory, accountancy, statistics, and taxation. Only the last will be here considered.

As already stated, no attempt will be made to cover the whole subject of income taxation, much less that of taxation in general. There will be no discussion of the merits of "progressive" or "regressive" income taxes. Nor will there be an exhaustive list of pros and cons regarding the income-tax base here proposed. The sole purpose will be to show how John Smith's income can be correctly ascertained for income-tax purposes.

The late Professor Thomas S. Adams had been much interested in *The Nature of Capital and Income* and told me he hoped it might be applied to taxation. But it was not until 1936, after his death, that I learned from Ogden L. Mills, former Secretary of the Treasury, that the bill for a "Spendings Tax"[21] introduced by him on July 20, 1921, when a member of Congress, had been partly suggested by Professor Adams.

This excellent and excellently drawn bill (H.R. 7867) seems to me to come nearer to a true income-tax proposal than any other bill hitherto introduced, and to fit, almost perfectly, the theory here developed. It aims to tax what is here called income *spent*, M_s.

For a generation this problem of defining taxable income in a satisfactory way has cried out for solution. Few practical problems in economics are more important. The just or unjust disposition of billions of dollars depends on how we define income. It is common knowledge that our present laws and the decisions of the Supreme Court under those laws often give manifestly absurd results.

The best practical formula for income-tax purposes would appear to be Formula (29).

In this formula, M_r includes not only John Smith's dividends; interest and rents; but the proceeds of sales of stocks, bonds, real estate, and all other property; payments to him of debts, or partial payments on them; money received as loans; all gifts of money, all bequests in money; money taken out of savings banks; and all other monies, if any, received within the year, except in payment for work.

M_d includes not only all money disbursed by John Smith for running

[21] Ogden L. Mills, "The Spendings Tax," *Bulletin of the National Tax Association*, October, 1921, pp. 18–20. See also John B. Canning and E. G. Nelson, "Budget Balancing and Economic Stabilization," *American Economic Review*, Vol. 24, March, 1934, pp. 26–37.

Irving Fisher, "The Income Concept in the Light of Experience," English reprint. The original publication of this article was in German in Vol. III of the Wieser Festschrift, *Die Wirtschaftstheorie der Gegenwart*, Vienna, 1927.

a business, but all capital expenditures, all money paid out in purchase of stocks, bonds, or other property, all money in payment of debts to banks or other creditors, all money lent, all gifts made in money, all actual payments by him into sinking funds, depreciation funds, funds for obsolescence, depletion, repairs, replacements, betterments, all actual payments into reserves of all sorts.

We know that $M_r - M_d + M_w + m$ must be absolutely equal to $\sum M_s$, the money spent by John Smith on consumption goods or their services.

But to this M_s, or *cost* of consumption goods, is now to be added a correction c for the use-cost discrepancy, the calculation of which correction has already been described in detail.

Since the present object is to ascertain the money value of John Smith's real income, and since it should be possible to make an official checkup by comparing the two equivalent values of I_{ba} given in Formula (29) by examining Smith's $\sum M_s$ (as well as another checkup through $\sum M_s = \sum M_{cr} - \sum M_d$), it is important that the constituent parts of $\sum M_s$ should be capable of exact specification.

For workmen, M_s has been thoroughly itemized in the studies of workmen's budgets, especially those of the United States Bureau of Labor Statistics.

The itemization for the well-to-do (who are those now paying income taxes) has never been as thoroughly done. There have been a few studies in intermediate budgets, including some by Dr. Royal Meeker and by Professor Jessica B. Peixotto of the University of California.

All of these studies, including workmen's budgets, show substantially the same categories constituting consumption goods.

It would take too much space to list here all the sub-categories. But all the important classifications, either for rich or poor, are included in the following summary groupings.

ITEMS OF BUDGET

1. Food
2. Clothing
3. Shelter
   (including installments and interest; taxes and assessments; repairs; water rent; fire insurance—on house)
4. House Operation
   (including house furnishings; fuel and light; other—e.g., ice, telephone, service, garbage removal, housecleaning supplies, laundry sent out, fire insurance—on furniture)
5. Amusements
   (including theatre, moving pictures, concerts, games, vacation, automobile, tobacco, club dues, social entertainment, other recreation)
6. Education and Church
7. Life Insurance
8. Contributions
   (including dependents, charity, gifts)
9. Miscellaneous
   (including medical care, barber, cosmetics)

All of these items represent components of Smith's taxable real income except charity, dependents, gifts, taxes and assessments, and life insurance, which should be and usually are, as a matter of policy, exempt.

As to a life insurance premium, unlike fire insurance or any other insurance, it consists of two separable parts—one a payment for the current risk and the other an investment. In respect to the latter, the life insurance company serves as a savings bank and the accumulation of the investment portions of the premiums gives John Smith a capital which he can draw upon at any time. Evidently this savings portion of the premium should not be counted or taxed as current income. It, like any other savings, is a carry-over for the future. The other portion —the payment for risk of death—is a sort of gift to Smith's descendents, paid in installments; and it should be treated like "charity, dependents, gifts." If these be tax-exempt, the whole of the life insurance premiums must be excluded from taxable income.

The foregoing items constitute $\sum M_s$, Smith's taxable income except for the correction c. But because these items are so numerous they are not well suited for an income-tax return; whereas, their exact equivalent (consisting of far fewer individual items) is well suited; namely $(\sum M_r - \sum M_d) + \sum M_w + m + c$.

The proposed Income Tax Return would then consist substantially of the following items, each of which is preceded by the formula letters in order to help the reader identify them.

TAX SCHEDULE TO BE FILLED OUT BY TAXPAYER

1.	M_w	Money from Salaries, Wages, Fees, Commissions .
2.	$M_r - M_d$	Money from Dividends . . . .
3.	$M_r - M_d$	Money from Rents and Royalties . . . .
4.	$M_r - M_d$	Money from Private Business or Profession, Partnership, Syndicates, Pools, less money put in during the year . . . .
5.	$M_r - M_d$	Money from Money borrowed less money lent or paid on loans . . . .
6.	$M_r - M_d$	Money from Interest received less interest paid . .
7.	$M_r - M_d$	Money from Sales of securities or other property less purchases (including all investments made in the year) . . . .
8.	$M_r - M_d$	Money from Gifts, Bonuses, Bequests . . . . . .
9.	m	Money from Use (depletion) of Cash . . . . . .
10.	c_h	Computed value received from use of residence less money spent on it (detailed in separate schedule 10) . . . .
11.	c_a	Computed value received from use of automobile less money spent on it (detailed in separate schedule 11) . . . .

12. c_f Computed value received from use of house furnishings less money spent on them (detailed in separate schedule 12)

13. $c = c_h + c_a + c_f$ (Sum of lines 10, 11, 12)

14. $M_s + c = \sum M_w + \sum M_r - \sum M_d + m + c =$ Total of lines 1 to 12

Deductions

15. Contributions

16. Taxes Paid

17. Life Insurance

18. Other deductions authorized by law

19. Total deductions (Sum of lines 15, 16, 17, 18)

20. Line 14 less line 19

Schedule 10 (detail for line 10)

Computed value received from use of residence (if owned or in process of being purchased) less money spent on it

(If living in rented home enter "zero" in line 10)

21. C_h Original cost of residence (with date acquired)

22. h 10%

23. h' Fraction of year owned (From ____ to ____)

24. $h'hC_h$ Computed value of use of residence

25. M_h Payments for residence (purchase price, part payment, interest or principal on mortgage, cost of repairs, betterments, taxes, insurance—and all other costs, if any, incurred by virtue of owning the home, but not including cost of running or care of the residence, as these items are otherwise included, nor duplicating any items already accounted for in lines 5, 6, 7, 8)

26. c_h Excess of 24 over 25 (for entry on line 10)

Schedule 11 (detail for line 11)

Computed value received from use (if owned or in process of being purchased) of automobiles (and motorboats, yachts, airplanes, etc.), less money spent on them

27. C_a Original cost of automobile (with date acquired)

28. a 20%

29. a' Fraction of year owned

30. $a'aC_a$ Computed value of use of automobile

31. M_a Payments for automobile (purchase price, part payment, interest or principal on mortgage, cost of repairs, taxes, insurance, and all other costs, if any, incurred by virtue of owning the automobile but not including cost of running or care of the automobile, as these are otherwise included, nor duplicating any items already accounted for in lines 5, 6, 7, 8)

32. c_a Excess of 30 over 31 (for entry on line 11)

Schedule 12 (detail for line 12)

Value received from use of house furnishings (if owned or in process of being purchased) less money spent on them

(If living in rented *furnished* house enter "zero" in line 12)

33.	C_f	Estimate of value of furniture . . . .
34.	f	15% . . . .
35.	f'	Fraction of year during which the bulk of house furnishings used was owned . . . .
36.	$f'fC_f$	Computed value of use of house furnishings . . . .
37.	M_f	Payments for house furnishings (purchase price, part payment, cost of repairs and replacements, taxes, insurance, and all other costs, if any, incurred by virtue of owning the furniture, but not including cost of care, as these are otherwise included, nor duplicating any items already accounted for in line 5, 6, 7, 8) . . . .
38.	c_f	Excess of 36 over 37 (for entry on line 12) . . . .

The above form of tax return is similar to those now used for the United States Income Tax except for the following items, excluded and included:

A. The money items to be excluded from the proposed schedule but which are entered in the United States tax schedules are: "capital gain and loss" (although all sales and purchases *within the year* are included in the proposed form), "losses by fire, storm, etc.," and "bad debts." Moreover, the 4th item in the proposed schedule, "from private business or profession, partnership, syndicates, pools, less money put in" differs from the corresponding item now usually entered "income (or loss) from partnerships, syndicates, pools, etc." by excluding any allowance for "bad debts" or "depreciation, obsolescence and depletion," and by excluding the two inventory items (for the beginning and end of the year). All these are essentially items of capital gain or loss. Only the actual net money taken out of the business would be used in the proposed income-tax schedules (though all the above items are extremely useful in the bookkeeping of the business itself).

In the same way only the actual money received from fiduciaries is included, with no bookkeeping adjustments.

B. The money items to be included in the proposed schedule but not entered in existing tax schedules are: money borrowed and lent, money disbursed for investments, gifts, bonuses, bequests, and the net cash used (m).

C. The item c is included in the proposed schedule but not properly accounted for in existing schedules.

The extra items included largely take the place of those excluded. They yield, with less bookkeeping and less error, substantially the same

result except for: (1) certain capital gains or losses accrued *during the year*, such as undivided profits, now improperly included; (2) the irrelevant items now improperly included from *other* years but entered under realized "capital gain or loss" of the current year; (3) sales and purchases, now imperfectly included; (4) real income other than represented by spendings or spendings not represented in the real income of the year as spendings for a house or furniture.

It has been shown that in computing income we must not add or subtract any mere changes in capital valuations, G. Just as a *rente*'s capital value may go up and down without increasing or decreasing the income, so capital gain or loss of any kind may be great or small without affecting the income of the year. Only the actual M's, monies taken in or paid out, are to be counted and counted only in the year in which these payments occur. Capital gain merely symbolizes future income. When that income comes will be the time to tax it. John Smith should not be required to pay more tax this year because of more income next year.

And, contrariwise, no allowance should be made for losses from shrinkage in value of property, even when it shrinks to zero, by fire, flood, drought, storm, shipwreck, earthquake, or anything else. Such loss merely symbolizes lack of future income. If John Smith's income in future years will be lessened because his factory burns up this year, he will then—in those future years—have correspondingly less taxes to pay. He should not have any less to pay *this* year on account of less income *next* year.

When a salaried man pays a tax this year on his whole salary, including what he saves and invests, and in future years pays taxes on the income then coming in from this year's savings, he is the victim of double taxation. That is, he is taxed on savings, S, this year and thereafter on Si every year (assuming that S yields income at the rate of interest i perpetually). If the tax rate is r, his tax this year is rS and every year afterward rSi, the present value of which is rS. Thus $2rS$ is, in effect, his tax burden because he could "compound" for his future perpetual tax rSi by a lump sum payment of its capital value, namely rS.

This is illustrated in the *Nature of Capital and Income*. Numerical examples are given by supposing three brothers each inheriting the same fortune, $100,000, but investing it differently, although all at 5%. One receives and spends $5,000 annually and is taxed 1% thereon. His total tax burden measured by the present value of his tax of $50 a year is $1,000.

The second accumulates his $100,000 at compound interest for 14 years, spending nothing. He then has $200,000 and thereafter receives

and spends $10,000 a year. But if he is taxed 1% not only on this but on his capital gains in the previous 14 years, he bears a total tax burden, measured in present value, of $1,714. The extra $714 represents income taxed twice.

The third brother spends $20,000 a year which exhausts his capital with interest in 6 years. If he is taxed only on his interest his tax burden is measured in present value by $158. The deficiency below $1,000 represents income not taxed at all.

But if the three taxes were adjusted according to the principles of this study the three tax burdens would all be equal, each being $1,000.

The present capital-gain tax is a haphazard tax, occasionally just, usually unjust, and seldom self-consistent. It does not even tax capital gain as it steadily accrues. It is not a true capital-gain tax nor a true income tax.

The present capital-gain tax is worse than haphazard. For it is subject to manipulation by clever tax-evaders or tax-avoiders. The taxpayer often has it in his power to choose when, if at all, he will make the sale and take the capital gain, and to choose what to sell. Naturally, he will so choose as to make the realized capital gain as small as possible and even negative if possible—a capital loss. That is, the tax dodger has an impulse to sell after a fall of price, thus increasing the fall, and a reluctance to sell after a rise, thus increasing the rise.

Consequently, among other faults of the present system, we find that it intensifies booms and intensifies depressions—tends to "boom the booms and bust the busts." It was doubtless one of the contributory causes of the stock market boom of 1929 and the depression following.[22]

Again, the loss taking reduces the taxes at the very time that the government most needs income.

Another absurdity, flowing directly from the attempt microscopically to calculate the income involved in a capital gain from a particular article of capital, is the solemn procedures supposed to "identify" any particular share of stock sold. When was it bought? If the luckless taxpayer, by inadvertence, happened to part with a certificate which had been acquired below the present price instead of one just like it but acquired above that price, he must report a gain instead of the loss which he might have taken. Obviously, it should make no difference whether one certificate or the other is sold. They represent legally exactly the same rights. One might as well tax all red certificates but no blue ones. To "identify" a certificate of stock is often difficult, sometimes impossible, and should never be necessary any more than identi-

[22] See *Booms and Depressions*, Adelphi Co., 1932.

fying each dollar we pay out from our check book, as some particular dollar acquired at a particular date.

Again it is common sense that a man who sells and immediately repurchases a stock should not be considered as having on that account a larger or smaller income than a man who merely holds the same stock.

Our present system gives such bizarre results that, instead of being carried out completely and consistently, it has been patched up by all sorts of special exceptions, regulations, rules, and conditions—such as that a capital loss can only be applied against capital gain and not against income in general, or that a stock sold and rebought within two weeks must be considered to affect one's income very differently from what it would be if one day more were to elapse.

In short, the offsetting of a receipt in one year by an outlay in another year is a delusion and a snare. No such pairing is necessary or desirable. All the proper offsetting is automatically taken care of by considering every sale within the year a plus and every purchase within that same year a minus, irrespective of what is sold or what is bought.

If some "Will Rogers" among economists would arise to poke fun at the workings of the American capital-gain-as-income tax, he might laugh it out of existence. The first publication of large American "incomes" revealed the surprising fact that a neighbor of moderate means had in 1930 an "income" of $331,000! This high figure was due to distress selling of old family property. The true income, ordinarily about $10,000, had been reduced. Another person with about the same income according to any rational measurement was able to report in 1935 an "income" of *minus* $235,000!

The proposed method of computing John Smith's income is not only far more accurate than any other method now in use, but far simpler, and requires fewer items, especially fewer *appraised* items. It therefore offers far less chance of evasion.

One reason for this is that two checkups are available on the money items, which usually constitute the major contribution to the total income. As we know, the income is to be calculated as $\sum M_r - \sum M_d + M_w + m + c$. But this is necessarily equal to $\sum M_s + m + c$, and also to $\sum M_{cr} - \sum M_d + m + c$. John Smith's books, therefore, afford the tax inspector a double check. The three figures must check to the last cent. There is no corresponding check available under the present system.

Another important guarantee of greater accuracy and freedom from evasion is that the proposed system does not much depend on records prior to the current year. Ordinarily, the only items which depend on earlier records are the original cost of an owned home and automobile.

The elimination of all mere bookkeeping entries for depreciation, for reserves, and for other devices to spread non-recurring expenses over

other years helps enormously in simplicity, ease of calculation and accuracy. In the proposed system almost all these calculations, which now must be so laboriously worked out for each capital item separately, take care of themselves. Their net result is, in effect, lumped together in the three corrections for house, household effects, and automobile. Only these, as has been noted, require the accountant's smoothing art. As to the rest, if John Smith conforms to approved norms, any sum in one item, too large, or too small, will be offset in other items, too small or too large. That is, the more items included, the less need of devices for smoothing or spreading. In fact, it is doubtful if even the three special cases—house, house furnishings, and automobiles—really need very often to be smoothed. The use-cost discrepancies will usually be almost negligible in their net effect. The reason is that installment buying, mortgaging, mortgage liquidation, and other devices are commonly used to such an extent that the individual will seldom be paying, in $\sum M_s$, much more or much less than the actual house use, automobile use, and house-furnishings use. This is no accident. Every normal individual, confronted with a large non-recurring item, whether a receipt or a disbursement, will seek ways of spreading it over time.

If John Smith does not do this, the results of his veering toward the spendthrift type or toward the miser type will automatically appear in the final figures. The result will be (for the current year) more taxes in the former case and less in the latter. The opposite is the case under the United States income tax as it now stands.

The proposed system also makes it impossible to evade, as rich men have so often done, by setting up a "personal company" and "borrowing" of it their living expenses. The tax evader is thus really getting income but calling it a "capital item" and therefore not income under the law. The accountant's ingenious devices for distinguishing income from capital merely afford the tax evader his opportunity. This fact is enough to condemn these distinctions as fallaciously made. The true distinction between income and capital is simple and clear enough, and permits of no evasion. This is the distinction between services and their capitalization.

It is not meant to argue against the legitimacy of ever taxing capital. In an emergency, as in a great war, we sometimes resort to capital levies. But whether or not, or when, taxes on capital are legitimate, we must always recognize that taxing present capital is taxing future income in advance. It means double taxation. This may be excusable in a great emergency, although, even then, the rate of the capital levy would presumably be much less than the rate on its future income. But the present taxes are the product not of an intelligent understanding of

how to tax for some great emergency, but of a confusion between capital and income.

It should also be noted that, if capital gain is to be taxed, corporations which are largely catch basins for capital accumulations would be taxed also. So would savings-bank accounts.

Finally, it should also be noted that the restriction to "realized" capital gain is impracticable—in fact, usually meaningless if confined to the current year, since not always will both the purchase and sale fall within the same year. Thus, even if capital, capital gains, or savings of any sort are to be taxed, the restriction to "realized" items should be out of the question. The distinction between "realized" and "unrealized" capital gain can only play havoc with our calculations and continue to make our results haphazard instead of systematic.

We may insert a word here on the somewhat analogous subject of stock dividends.[23]

The Supreme Court long ago decided that, in general, stock dividends are not taxable because the receiver of them has no different ownership from what he had before he received them, though that ownership is represented by more pieces of paper. This is, of course, true. But it does not touch the real problem, which is whether John Smith's stock dividend leads or does not lead the way to more real income for him.

If the stock dividend is 100%, that is, is a "split-up" of two shares for one, the chances are there will be no resulting real income. But if the dividend is 1%, there is a good chance that Smith will sell the new stock and spend the proceeds on living expenses, so that, so far as he is concerned, he will get exactly the same income as if he had received his dividend in money. A vague feeling that small stock dividends enable a person to evade the income tax is behind the opposition so often voiced to the Supreme Court decision referred to.

Under the *proposed* system, however, if the proceeds go for living expenses, there will be a tax; but if the stock is sold and proceeds reinvested in some other stock, no tax will result. Also if it is not sold at all, there will be no tax. All three of these results are as they should be. Under the *present* law, on the other hand, if the stock is sold and the proceeds reinvested in another stock, Smith is taxed, though he should *not* be.

[23] See Fred Rogers Fairchild, "The Economic Nature of the Stock Dividend," *Bulletin of the National Tax Association*, April, 1918, pp. 161–163; also June, 1918, pp. 240–243; Henry H. Bond, "A Practical Aspect of the Stock Dividend Question," *Bulletin of the National Tax Association*, June, 1918, pp. 237–240; Edwin R. A. Seligman, "Are Stock Dividends Income?" *American Economic Review*, September, 1919, pp. 517–536.

On May 16, 1936 a new decision of the Supreme Court made a new distinction. If a stock dividend of common stock is given to preferred stockholders, it is taxable because the preferred stockholders are then supposed to get a real addition to their property.

Five months later, as might be expected, the Board of Tax Appeals, in trying to apply the Court decisions, concluded that "the matter has not been, so far as we know, squarely decided in any of the cases." It should be clear to the reader that, under the method of levying an income tax here proposed, there could be no ambiguity.

Since, as has been noted, corporations and other artificial persons can have no income separate from, and in addition to, the income of their shareholders, bondholders, and employees, they should not be taxed under an income-tax law unless, as mentioned above, merely as agents for these real persons on the principle of stoppage at the source.[24] Any taxes on corporations *in addition* to those properly allocated to real persons means doubly taxing the common stockholders.

The prevalent idea that corporations should be subject to an income tax independently of real persons is largely based on the confusion already so often noted between income and earnings, which implies also the confusion between income and capital.

This very problem has recently been thrust before us by the tax on the undivided profits of corporations.

As originally framed by experts of the United States Treasury, this proposal to tax undivided profits included taxing them at varying rates up to nearly 50%. A tax of 50% would mean that, out of $100,000 undivided profits available for ploughing back into a corporation's business, the Government would take $50,000, leaving $50,000, the other half, to be actually ploughed back. Whatever this $50,000, belonging to the stockholder, might earn in the future, would also be subject to tax in the future. If the rate of this tax were also 50% perpetually, the stockholders would lose 50% of all these future earnings from that $50,000. But to take away half the earnings of this $50,000 is to take away all the earnings of half of this $50,000, which, if continued forever, is to deprive the stockholders of any benefits from said half.

Moreover the other half, $25,000, will, beginning a year later, be likewise split in two, one-half, or all its future benefits, going to the Government. Only the other half, $12,500 in present value, remains and half of that, after two years, would virtually belong to the Government. Evidently the stockholders will have left only half of half of half indefinitely, or zero. They will be merely holding property to earn taxes for the Government.

[24] Such a tax paid by a corporation on behalf of a stockholder may conceivably constitute an overpayment and so require, later, a refund by the government.

Of course, the object of such a tax, however, is to force the distribution of dividends, the desirability of which is a question which does not concern us here.

CHAPTER XII: TERMINOLOGY

It is my earnest hope that the foregoing presentation of the income problem will carry conviction. So far as I can see, the only possible controversy will be over terminology. Some persons, while admitting all my contentions as true under my concept of income may prefer still another income concept. And there may be some for whom such other income concept is felt to be a serious matter.

I might claim that it ought not to be a serious matter—that anyone has a right to define any word to suit himself. But this would be dodging the issue. For most people and for many economists words are more important than ideas.

There are two chief criteria for choosing a terminology. One is demanded by science, the other, by usage.

Briefly, the first is that the term shall denote a scientifically useful concept. To define income as anything received on weekdays but not on Sundays, or as receipts by those over twenty-one years of age, or (as an able economist once actually proposed) as any economic good used only once, would scarcely fulfill this first, or scientific, requirement. But the income concept here adopted, which I have advocated since 1897, certainly does fulfill this requirement and in several ways.

First and foremost, under this concept, capital and income are associated definitely, clearly, simply. For, under this concept, capital value is the capitalized or discounted value of future expected income. Under no other concept is this true. It is not true of earnings, as we have seen (unless the parts of earnings consisting of capital gain, before being discounted, be displaced from those positions in time where they belong).

Again, the chosen concept articulates all its numerous sub-concepts —the income from a particular instrument of capital, the income from any designated group of instruments, the income of any particular individual, the income of any designated group of individuals, the income of society, money income, income spent, real income, psychic income. No other concept accomplishes this articulation. For this reason many writers are forced to frame a number of separate and almost unrelated definitions of income—psychic, real, spent, money, corporate, individual.

The most discordant element in the medley of definitions is savings. If, for instance, money income includes savings, while psychic income does not, a discrepancy between money income and psychic income necessarily arises.

So much for the requirements of science, where the problem is to

find a term to fit a pre-existing concept. We now come to the requirements of usage, where the problem is, on the contrary, to find a concept to fit a pre-existing term. The dictates of usage are never very clear. For this reason some scientists would ignore dictionaries. But we cannot do this with impunity. No scientist can ride roughshod over usage and call black white. Nevertheless, the fact that usage is not clear gives him the right, in fact the duty, to clarify. This question of interpreting the popular usage as to the term "income" has been argued elsewhere and in many places.[25] Extended discussion here would be out of place.

The disagreement over which many writers seem to feel most strongly is as to the inclusion or exclusion of capital gain or savings and I believe their feelings can be greatly assuaged by employing, when necessary, a qualifying phrase.

Just as accountants speak of income "before taxes" are taken out and income "after taxes" are taken out, so I now propose that, to avoid controversy, we speak of income "before savings" are taken out and income "after savings" are taken out; the latter evidently being what I call income proper.

In 1908 Professor Winthrop M. Daniels, criticizing my definition of income on the score of violating common usage, says that Dr. Johnson, "certainly a competent lexicographer," advised Boswell: "Live within your income. Always have something saved at the end of the year."

Such criticism of my use of the term "income" because it excludes savings, while Dr. Johnson seems to include savings, would obviously lose all force if the term "income" as used by me were replaced by the term "income after savings are taken out"; and if the term "income" as used by my critics were replaced by the phrase "income before savings are taken out."

Without this qualifying clause, or some equivalent, it cannot be denied that my income concept sometimes collides with popular usage. But so does every other income concept; for usage is inconsistent. Even as to including or excluding savings, usage is inconsistent. Thus corporate savings are not included by a stockholder in counting up his income but only the dividends he actually receives. Yet logically or scientifically it should make no difference as to savings being income whether the savings are made for the stockholder by the corporation or by the stockholder for himself as might have been done. Such inconsistency of usage would disappear if the proposed qualifying phrases were used. For income "after savings" would exclude all savings alike, whether made by the corporation or by the individual, and likewise income "before savings" would include all savings alike.

[25] For an assortment of definitions of income see my *Nature of Capital and Income*, 1906, Appendix to Chapter VII.

Let us see what the two new phrases, here proposed, "before savings" and "after savings," really mean.

As must have been noted, the process of saving consists of disservices, certain negative terms in the net algebraic service-sum constituting John Smith's income as defined in this study. Thus, when he "saves" $1,000, i.e., puts aside $1,000 in a cashbox, or stocking, or in a savings bank, or in the purchase of a bond or stock, there is thereby created a negative item in the accounting of John Smith's service-sum. We debit $1,000 to the cash box contents, stocking contents, savings bank account, bond or stock, or other investment.

But, while all savings are debit-items, not all debit-items are savings. Which items are to be considered savings depends, as has been noted, on what is the norm of reference, which may be arbitrarily adopted. But, assuming an adopted norm, the amount of savings debit is clear. Subtracting this debit, we get John Smith's net "income" or net service-sum. To call this net sum John Smith's "income after savings" is, under our definition, correct enough, though evidently tautological under the service-sum concept. It simply represents the *complete* sum, all terms, positive and negative, being included.

On the other hand his "income before savings" is not tautological; for then the list of terms is *incomplete*. The phrase thus designated is a *gross* income figure, one subject to certain deductions for reinvestment.

Thus income "before savings" includes more than this year's total net service-sum. What it includes is the capitalization of certain services to be received in *future* years. Savings are a present accounting for *postponed* services, services subtracted from the services of the current year and added (with interest) to the services of later years.

Savings are capital, the anticipation of income, not its realization. To add this year's anticipation of future years' services to this year's services actually realized is counting chickens before they're hatched, and therefore counting them twice.

We avoid all this when we tear away the money wrapper and look at the bundle of services which that money wrapper represents. It is significant that, while usage often sanctions the inclusion of savings in *money* income, it seldom includes any savings in *real* income; and never includes any savings in psychic income. But to include savings in some of these cases and not in others upsets, of course, the income articulations. This brings us back to the "scientific" criterion which primarily concerns us here. And no other income concept can qualify as providing for either capitalization or articulation.

IRVING FISHER

Yale University
New Haven, Conn.

METHODS OF ELIMINATING THE INFLUENCE OF SEVERAL GROUPS OF FACTORS

By Corrado Gini

The quantitative treatment of collective phenomena admits of several stages, or steps, all of which, even in the most refined statistical research, are but seldom carried out completely.

The first of these stages is the *quantitative description of the collective phenomena*. It is well known that one of the groups of researches from which statistical science draws its origin was confined precisely to this stage. To describe the *res notatu digna* of the State was the purpose of the course initiated by Hermann Conring in 1660 at the University of Helmstädt and generally the purpose of the so called "Universitary Statistics" which then flourished in Germany. Some of the statistical researches of Graunt and of other "Political Arithmeticians," which constituted the other source of modern statistics, were also purely descriptive in character, as when Graunt ascertained the ratio between the two sexes at birth in London and in the country parishes; but in other subjects their researches went a step further, looking for the causes of the quantitative differences of the collective phenomena that they had described. It is this aim which justifies the denomination of "Investigative Statistics" which has been applied to their school, in contrast to that of "Descriptive Statistics" applied to the German school of Conring and Achenwall.[1]

In such *explanation of quantitative differences in collective phenomena*, the quantitative element may, however, vary in importance.

Often the statistician is satisfied with indicating the circumstances (or some circumstances) which may explain the observed differences, founding his explanation on the influence which, on the basis of current experience or of theoretical considerations, may be attributed to these circumstances. So the higher death-rate which had been observed in the cities in comparison with the country was attributed to the more frequent recourse to artificial suckling, to the greater danger of the urban occupations, to the greater diffusion and gravity of vice, to the crowding of the people in the slum quarters, and to other hygienic shortcomings of the great urban agglomerations. In such explanations the quantitative element is completely lacking. The description of the collective phenomenon is quantitative, but the investigation for their explanation is qualitative.

The quantitative element is introduced into the explanation also, when we demonstrate on a statistical or experimental basis that the circumstances considered really have an influence in the direction stated. Thus, we may prove, on a statistical basis, that the infants

[1] Cf. our *Curso da Estadistica*, Barcelone, Editorial Labor, 1935, pp. 4 ff.

suckled at breast present a lower mortality than those artificially nourished, that the occupations prevailing in the cities present a higher death-rate than those prevailing in the country, that the death-rate increases with the crowding of the houses, and is higher among the vicious than it is among ordinary people. Generally, the quantitative treatment of collective phenomena stops here; moreover, the influence of some circumstances is sometimes so obvious (as, for example, the influence of vice on the death-rate) that it is understandable that a quantitative demonstration of it seems superfluous.

Even more seldom does it appear necessary to give a quantitative proof that the two groups compared are really differentiated according to the circumstances considered. Such necessity is felt only in some doubtful cases; even without any statistical evidence we are prepared to admit, for example, that artificial nourishing of infants or that adult vices are more common in the city than in the country.

This is true, however, only in so far as we limit our aim to inquiring *what are the causes* of the observed differences, which constitutes after all a qualitative explanation, even though it requires some quantitative researches. In order to have a quantitative explanation, we must take a step further and inquire also *what is the bearing of these various causes* in determining these differences. After having recognized that artificial nourishing of infants, overcrowding, unhealthful occupations, and other unhygienic conditions, as well as vice, are the essential causes of the higher death-rate of the city, we may ask to what extent does each of these circumstances contribute to the differential death-rate of urban and rural populations. Only in that case will the quantitative treatment of the question be complete. The substitution of a quantitative for a qualitative explanation of the statistically observed phenomena is, in my opinion, the direction in which applied statistics may realize in the future its greatest progress. Only when this step is made, will Demography, Biometry, Econometrics, Psychometrics, etc. deserve the denomination of exact sciences.

The bearing of certain circumstances on the determination of a difference between the intensities that a phenomenon presents in two groups of observations depends on various characteristics of these circumstances or of the groups considered:

(1) The *differentiation* of the two groups in terms of the circumstances considered. Thus artificial nourishing of infants may explain the higher death-rate of the urban population only in as much as it is more frequent in cities than in the country.

(2) The *influence* of every circumstance considered on the differential phenomenon, such an influence depending on the *direction* and on

the *intensity* of its effects. Thus, the extent to which the artificial nourishing of infants contributes to the explanation of the differential death-rate in urban and rural populations will be much or little according to the amount by which artificial nourishing is found to increase the infantile death-rate.

(3) The *extent* of the application of the circumstances considered. Some circumstances exert their influence on all the groups combined (as unhealthy general conditions), others, only on certain age classes (as artificial nourishing of children or occupational risks), and their effects on the general death-rate are manifestly more or less important according to the percentage represented by the respective age classes on the whole population.

(4) The *composition of the effects* of the different circumstances. There are circumstances the combination of which causes an effect more intense than the sum of the effects realized by them separately. Vice has a certain effect in increasing the death-rate and an unhealthy occupation has its own effect in the same direction; but the coincidence of vice and of an unhealthy occupation may have an effect much more serious than the sum of their separate effects. The same may be said of general hygienic conditions of the family and of artificial nourishing of children; as a matter of fact, artificial nourishing has always a bad effect on the infant death-rate but this is disproportionately increased by an unhygienic environment. There are other instances in which two bad circumstances tend to correct each other, as syphilis and malaria: both increase the general death-rate of the population, but the prevalence of malaria tends to reduce the bad effects of syphilis, the high fever caused by an attack of malaria killing the syphilis bacilli or reducing their virulence.

It is clear that, in order to measure the contribution of the different circumstances in the determination of a certain phenomenon, it is necessary not only to know their effects when each of them acts separately, but also to know the manner in which their effects combine when they act simultaneously.

It may indeed be said that it is often possible to measure the contribution of the various circumstances without knowing the differentiation, the influence, and the extent of application of the given circumstances; but it is always necessary to have some knowledge, or to make some hypothesis, about the composition of their effects.

As a matter of fact, when we have at our disposal a sufficient set of observations (for example, the death-rate of the urban and rural populations in many countries and, at the same time, the indices, for the same countries, of the artificial nourishing of infants, of overcrowding, of unhealthy conditions, etc.) we may calculate the bearing of these

various circumstances on the differential death-rate, basing our calculation on the principle of concomitant variations. But the application of this principle, through multiple correlation or some other method, cannot lead us to the determination of the contribution that each circumstance makes to the resultant effect, until we have established how the separate effects of the given circumstances combine in their simultaneous action.

The study of the composition of effects is thus indispensable for a complete quantitative treatment of statistical phenomena.

The ways in which the effects of a system of causes may combine are theoretically infinite and actually most varied. Genetics presents examples of all the degrees of dominance of the effects of one circumstance over those of another circumstance, sometimes with a whole hierarchy of dominance among a series of several circumstances; it presents examples of complementary or, on the contrary, of inhibitory circumstances; of supplementary or cumulative ones; of intensifying or, on the contrary, of diluting ones; of indifference or, on the contrary, of attraction (or vice versa of repulsion) between certain combinations. Chemistry presents examples of combinations which take place only between definite quantities of certain bodies, the quantities in excess remaining unaffected; it, as well as physiology, presents examples of "all-or-none" reactions, that is, of effects which are produced only when certain circumstances attain a given limit, and are not increased when these limits are surpassed. In my opinion these sciences should suggest to the sociologist inspiring analogies for the analysis of the multifarious effects of the simultaneous action of various circumstances. Statistics, losing in variety what it aims to gain in precision, must be content with more simple schemes.

The simplest scheme is that which may be called the *summatory scheme*, according to which the total effect of several circumstances acting simultaneously is equal to the sum of the effects of the single circumstances acting separately. The method of residues, one of the four pillars of the inductive logic of John Stuart Mill, that the statisticians have introduced in their treatises, is based precisely on this hypothesis, as I indicated twenty-five years ago.[2] In some special cases the hypothesis is certainly justified and consequently the application of the method is correct. Thus, we may calculate the net immigration of a country by subtracting the natural increase (derived from the excess of births over deaths) from the total increase of population (calculated

[2] *Intorno al metodo dei residui dello Stuart Mill e alle sue applicazioni alle scienze biologiche e sociali.* Studi economico-giuridici della R. Università di Cagliari, 1910.

by the difference between the figures of two censuses). It is clear as a matter of fact that the total increase of population results from the algebraic sum of births, deaths, immigration, and emigration. But in many other cases the summatory scheme evidently does not fit the facts; so that it is surprising that for so long a time the hypothesis which is basic to it was not recognized. In particular, the said hypothesis is basic to the method of multiple correlation and, more generally, to the method of least squares (as it is usually applied to functions linear in respect to the parameters) and here also, as far as I know, the hypothesis has scarcely been noticed nor has its bearing been adequately investigated, as would be necessary for judging the reliability of the results.[3] To be fair towards statistics, we should recall that the same hypothesis is adopted also in applied mechanics, where it is indicated as the *principle of superposition of effects*. We must add, however, that here the principle is accepted only in so far as small forces are in action, a limitation which perhaps may be usefully introduced also in the applications of the method of multiple correlation.

Although the summatory scheme is basic to some general statistical methods, in special cases, in which it manifestly would not fit the facts, the statisticians base their procedure on another scheme. Thus the amount of the wealth of a country is the resultant of the individual fortunes and of the number of individuals, but nobody thinks of obtaining an index of the individual fortunes by *subtracting* the number of individuals from the amount of the national wealth; the index is obtained by *dividing* the amount of wealth by the number of individuals. In this case, the scheme which is appropriately applied is the *multiplicatory scheme*, according to which the total effect of several circumstances acting simultaneously is equal to the product of the effects of the single circumstances acting separately.

This scheme is also applicable when we obtain the coefficient of variability by dividing the mean deviation by the average of the character, as well as when we obtain the coefficient of correlation by dividing the coefficient of regression by the ratio of the standard deviations of the two correlated characters.

Moreover, we may remark that only the multiplicatory scheme justifies our calling the different circumstances which contribute to a given result *factors*.

On the multiplicatory scheme also is based the application of the so-called "ideal formula" to the computation of the index numbers of

[3] The only reference to this problem that I have been able to find is in Roos's *Dynamic Economics*, to which my attention has been called at the time of reading proofs. Cf. C. F. Roos, *Dynamic Economics*, The Principia Press, Bloomington, Indiana, 1934, pp. 47–48, 89 ff.

prices and to other instances of the methods of elimination. This is a subject which deserves special attention.

The variations presented by the average prices of a group of commodities from time to time or from one locality to another may be considered as the resultant of two contributing circumstances, the variations of prices for the single commodities and the variations of the relative importance of these commodities. The index number of prices measures the variations of the first component (prices) by eliminating the variations of the second component (relative importance of the different commodities). Similarly, the method of standard population measures the variations of the death-rate (or of the birth-rates, or of the marriage-rates, etc.) by eliminating the variations of the different age composition of the populations. We may, however, make different hypotheses concerning the manner in which the two components combine their effects. Some authors, for example, have subtracted from the general death-rates observed for one population, the corresponding death-rate calculated for the standard population, in order to measure the influence that the different age composition exerts on the general death-rate of the population considered. These authors have applied the summatory scheme. It may, however, be remarked that the number of deaths in any age class is obtained by multiplying the living persons by their death-rate, and similarly the value of a category of commodities is given by the product of price and quantity of the commodity. In every age or commodity category it is then the multiplicatory scheme which is applicable. It seems then justifiable to apply also to the average price of all commodities or to the general death-rate for all age classes, this scheme, rather than any other. The adoption of the multiplicatory scheme in the methods of elimination is in keeping with the adoption of the ideal formula, for, according to this, the index number of one component (for example, prices) multiplied by the index number of the other (quantities) gives the index number of the resultant phenomenon (values).

Until the present time, however, the methods of elimination through the ideal formula have generally been applied to the simplest case of phenomena resulting (or considered as resulting) from the combined effect of two factors (or groups of factors): prices and quantities of different commodities; death-rates and size of different age classes, and so on.[4] Often, on the contrary, we meet phenomena which may be con-

[4] So far as I know, the only exception is the note by J. K. Wisniewski: "Extension of Fisher's Formula Number 353 to Three or More Variables" *Journal of the American Statistical Association*, March 1931, pp. 62–65, to which my attention was called by Prof. Ragnar Frisch at the time of reading proofs.

sidered as the resultant of several factors. For such cases is it possible to indicate formulas based on the same scheme and having the same properties as the so-called ideal formula for the case of two factors? An answer to this question is the object of the following sections of this paper. We will consider the case of three factors, but the results may easily be generalized to any number of factors whatsoever.

We shall indicate the three factors by the letters a, b and c and their resultant effect by the letter T. The first object of our inquiry will be to indicate the formulas for an index of the variations of a by eliminating the variations of b and c.

The formulas will be different according to the different types of interrelations which exist among the three factors. It may be useful in this connection to distinguish three types of such interrelations.

Type A. The effect T is the resultant of the action of the factor a, on one hand, and of the joint action of the two *separate* factors, b and c, on the other. Example: The general death-rate may be considered as the resultant of the specific death-rates for every category of a given age and of a given sex, on one hand, and, on the other, of the composition of the population by age and by sex. The problem is that of constructing an index number of the variations of the specific death-rates by eliminating from the variations of the general death-rate the influence of the sex and age composition of the population.

Type B. The effect T is the resultant of the action of the factor a, on one hand, and, on the other, of the joint action of two factors, b and c, one of which is *supplementary* to the other. Example: The annual salary of an employee may be considered as the resultant of his salary per hour, on one hand, and, on the other, of the number of working days per year and of working hours per day. The problem is that of constructing an index number of the variations of the hourly salary by

---

The object of Wisniewski's note is the same as that of the following sections of this paper; but the theoretical treatment as well as the application can hardly be considered satisfactory. The case considered by Wisniewski is that in which the formula is symmetrical in respect to the three or more variables, and consequently corresponds to our Type C, while the application (value of some agricultural products depending upon prices, area harvested, and yield per area unit) falls within the scheme of our Type B. The formula arrived at (using the notation adopted in our Type C) is:

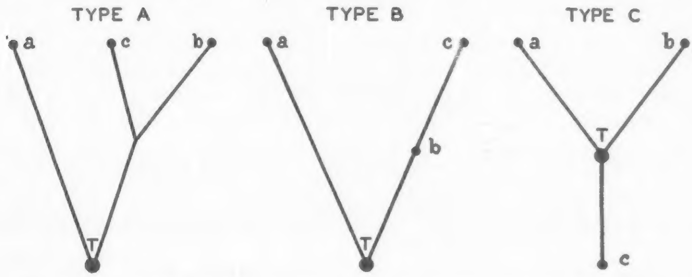
$$I_p = \sqrt[12]{\frac{(\sum p_{2i}q_{2i}r_{2i})^4}{(\sum p_{1i}q_{1i}r_{1i})^4} \cdot \frac{(\sum p_{2i}q_{1i}r_{1i})^2}{(\sum p_{1i}q_{2i}r_{2i})^2} \cdot \frac{\sum p_{2i}q_{1i}r_{2i}}{\sum p_{1i}q_{2i}r_{1i}} \cdot \frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{2i}}}$$

It is difficult to understand (and this seems to me the main objection) how this expression can be considered as an index of prices. As a matter of fact, the factors in the above expression are not averages of the individual price relatives p_{2i}/p_{1i}.

eliminating the influence of the variations of the number of working days per year and of the number of working hours per day. Such an index may be reduced to an index of the variations of the hourly salary by eliminating the variations of the number of working hours per year.

Type C. The effect T is the resultant of the *symmetrical* action of the three factors a, b and c. Example: The volume of the chest may be considered as the resultant of the product of the two diameters (transverse and antero-posterior) of the thorax and of the sternal height. The problem is that of constructing an index number of the variations of one of the three linear dimensions by eliminating the variations of the other two.

The following figure is a convenient illustration of the differences between the three types:



The second object of our inquiry is to compare the formulas obtained for the variations of a after the elimination of the variations of b and c with those obtained after a less complete elimination, that is, after the elimination only of the variations of one factor, for example, b.

In order to fix the ideas, we will have in view, in our treatment, the special examples described above.

TYPE A

We shall indicate by:

p_{1i}', the specific death-rate of the males of a certain age class i in population 1;

p_{1i}'', the specific death-rate of the females of this age class i in the same population 1;

p_{2i}', the specific death-rate of the males of the age class i in population 2;

p_{2i}'', the specific death-rate of the females of the age class i in the same population 2;

q_{1i}', the number of males included in the age class i in population 1;

q_{1i}'', the number of females included in the age class i in the same population 1;

q_{2i}', the number of males included in the age class i in population 2;

q_{2i}'', the number of females included in the age class i in the same population 2;

$q_{1i}=q_{1i}'+q_{1i}''$, the number of individuals (males and females) included in the age class i in population 1;

$q_{2i}=q_{2i}'+q_{2i}''$, the number of individuals (males and females) included in the age class i in population 2;

p_{1i}, the specific death-rate of the q_{1i} individuals included in the age class i in population 1;

p_{2i}, the specific death-rate of the q_{2i} individuals included in the age class i in population 2.

Thus

$$p_{1i} = \frac{p_{1i}'q_{1i}' + p_{1i}''q_{1i}''}{q_{1i}}, \tag{1}$$

and

$$p_{2i} = \frac{p_{2i}'q_{2i}' + p_{2i}''q_{2i}''}{q_{2i}}. \tag{2}$$

The sum $\sum p_{1i}q_{1i}=\sum p_{1i}'q_{1i}'+\sum p_{1i}''q_{1i}''$ represents the total number of deaths, and the ratio $\dfrac{\sum p_{1i}q_{1i}}{\sum q_{1i}}$ describes the general death-rate of population 1.

Similarly the sum $\sum p_{2i}q_{2i}=\sum p_{2i}'q_{2i}'+\sum p_{2i}''q_{2i}''$ represents the total number of deaths, and the ratio $\dfrac{\sum p_{2i}q_{2i}}{\sum q_{2i}}$ describes the general death-rate of population 2.

The ratio

$$I = \frac{\dfrac{\sum p_{2i}q_{2i}}{\sum q_{2i}}}{\dfrac{\sum p_{1i}q_{1i}}{\sum q_{1i}}},$$

gives the index number of the variation of the general death-rate from population 1 to population 2.

Now what part of this variation depends upon the variations of the age composition of population and what part upon the variations of the specific death-rates for the single age classes?

Applying the ideal formula, the index number of the variations of the specific death-rates for the single age classes is

(3) $$I_p = \sqrt{\frac{\sum p_{2i}q_{1i}}{\sum p_{1i}q_{1i}} \frac{\sum p_{2i}q_{2i}}{\sum p_{1i}q_{2i}}}.$$

We may say, in other words, that I_p is the index number of the variation of the general death-rate due to the variations of the specific death-rates for the single age classes or, again, that it is the index number of the variation of the general death-rate after the influence of the variations of the age composition of the population have been eliminated.

The reciprocal index number of the variation of the general death-rate due to the variation of the age composition of the population is

$$I_q = \frac{\sum q_{1i}}{\sum q_{2i}} \sqrt{\frac{\sum q_{2i}p_{1i}}{\sum q_{1i}p_{1i}} \frac{\sum q_{2i}p_{2i}}{\sum q_{1i}p_{2i}}}.$$

Now the problem is one of constructing, instead of I_p, an index number (which may be indicated as $I_{p',p''}$) that measures the variation of the general death-rate after we have eliminated the influence of the variations of the composition of the population both by age and by sex, or, in other words, an index number of the variations of the specific death-rates distinguished according to the single age classes and, at the same time, according to sex. The procedure is not different from that of the preceding case and leads to the following expression for this index number:

(4) $$I_{p',p''} = \sqrt{\frac{\sum p_{2i}'q_{1i}' + \sum p_{2i}''q_{1i}''}{\sum p_{1i}'q_{1i}' + \sum p_{1i}''q_{1i}''} \frac{\sum p_{2i}'q_{2i}' + \sum p_{2i}''q_{2i}''}{\sum p_{1i}'q_{2i}' + \sum p_{1i}''q_{2i}''}},$$

while the reciprocal index number,

$$I_{q',q''} = \frac{\sum q_{1i}}{\sum q_{2i}} \sqrt{\frac{\sum q_{2i}'p_{1i}' + \sum q_{2i}''p_{1i}''}{\sum q_{1i}'p_{1i}' + \sum q_{1i}''p_{1i}''} \frac{\sum q_{2i}'p_{2i}' + \sum q_{2i}''p_{2i}''}{\sum q_{1i}'p_{2i}' + \sum q_{1i}''p_{2i}''}},$$

measures the variation of the general death-rate due to the variations in the age and sex composition of the population.

Recalling (1) and (2), we may also write formula (3) in the following fashion:

(3 bis) $$I_p = \sqrt{\frac{\sum p_{2i}q_{2i}}{\sum p_{1i}q_{1i}} \frac{\sum \left[\frac{p_{2i}'q_{2i}' + p_{2i}''q_{2i}''}{q_{2i}}\right] q_{1i}}{\sum \left[\frac{p_{1i}'q_{1i}' + p_{1i}''q_{1i}''}{q_{1i}}\right] q_{2i}}},$$

while formula (4) may be written as follows:

$$\text{(4 bis)} \qquad I_{p',p''} = \sqrt{\frac{\sum p_{2i}q_{2i}}{\sum p_{1i}q_{1i}} \cdot \frac{\sum \left[\dfrac{p_{2i}'q_{1i}' + p_{2i}''q_{1i}''}{q_{1i}}\right] q_{1i}}{\sum \left[\dfrac{p_{1i}'q_{2i}' + p_{1i}''q_{2i}''}{q_{2i}}\right] q_{2i}}}.$$

A comparison of formulas (3 bis) and (4 bis) shows that they differ only in the members between square brackets, which represent weighted averages of the terms p_{2i}', p_{2i}'' in the numerator and of the terms p_{1i}', p_{1i}'' in the denominator. In formula (3 bis) the terms p_{1i}', p_{1i}'' are weighted according to the quantities q_{1i}', q_{1i}'' and the terms p_{2i}', p_{2i}'' according to the quantities q_{2i}', q_{2i}'' but in formula (4 bis) on the contrary the sums p_{1i}', p_{1i}'' are weighted according to the quantities q_{2i}', q_{2i}'', and the terms p_{2i}', p_{2i}'' according to the quantities q_{1i}', q_{1i}''. If $q_{1i}'/q_{1i} = q_{2i}'/q_{2i}$ (and consequently $q_{1i}''/q_{1i} = q_{2i}''/q_{2i}$), then $I_p = I_{p',p''}$.

This means that, if, in every age class, the sex proportion does not differ between population 1 and 2, the index number obtained by eliminating the influence of the age and sex composition of the population coincides with the index number obtained by eliminating the influence of the composition of the population only by age. This happens because, in such cases, there is no variation in the composition of population by sex, to influence the index number of the general death-rate after the influence of the age composition has been eliminated. If the variation of the weights q_i from the population 1 to the population 2 is correlated with the differences $p_i' - p_i''$, so that we may expect $\frac{q_{2i}'}{q_{2i}} > \frac{q_{1i}'}{q_{1i}}$ (or on the contrary $\frac{q_{2i}'}{q_{2i}} < \frac{q_{1i}'}{q_{1i}}$) when $p_{1i}' > p_{1i}''$ or $p_{2i}' > p_{2i}''$, then we have $I_{p',p''} < I_p$ (or, in the contrary case, $I_{p',p''} > I_p$).

This means that if the proportion of males tends to increase—or on the contrary to decrease—from the population 1 to the population 2 in the age classes in which the specific death-rate is higher for the male sex, the index number obtained by eliminating the influence of the age and sex composition is lower—or higher, respectively—than the index number obtained by eliminating only the influence of age composition.

We have considered an example in which the factor c (in our case the sex) presents only two modalities (male, female), but it is easy to generalize the formulas and the conclusions for the case in which the modalities of the factor c are more numerous.

TYPE B

We shall indicate by:

π_{1i}, the salary per hour of the employee i in year 1;

q_{1i}, the number of days worked by this employee i in the same year 1;

μ_{1i}, the number of hours per day worked by employee i in the same year 1;
π_{2i}, the salary per hour of the same employee i in year 2;
q_{2i}, the number of days worked by this employee i in year 2;
μ_{2i}, the number of hours per day worked by employee i in year 2;
$p_{1i}=\pi_{1i}\mu_{1i}$, the salary per day of employee i in year 1;
$p_{2i}=\pi_{2i}\mu_{2i}$, the salary per day of employee i in year 2.

The sum $\sum p_{1i}q_{1i}=\sum \pi_{1i}q_{1i}\mu_{1i}$ represents the total of the salaries paid in year 1 and similarly the sum $\sum p_{2i}q_{2i}=\sum \pi_{2i}q_{2i}\mu_{2i}$ represents the total of the salaries paid in year 2.

The ratio,

$$I = \frac{\sum p_{2i}q_{2i}}{\sum p_{1i}q_{1i}},$$

describes the index number of the variation of the total of salaries from year 1 to year 2.

The index number

$$I_p = \sqrt{\frac{\sum p_{2i}q_{1i}}{\sum p_{1i}q_{1i}} \frac{\sum p_{2i}q_{2i}}{\sum p_{1i}q_{2i}}}, \tag{1}$$

constructed according to the ideal formula, measures the variation of the total of the salaries which has occurred because of the variations of the individual daily salaries after the influence of the variations which occurred in the number of working days has been eliminated.

The reciprocal index number,

$$I_q = \sqrt{\frac{\sum q_{2i}p_{1i}}{\sum q_{1i}p_{1i}} \frac{\sum q_{2i}p_{2i}}{\sum q_{1i}p_{2i}}},$$

provides, on the other hand, the measure of the variation of the total of the salaries which is due to the variations in the number of working days.

Now the problem is one of constructing, instead of the index number I_p, an index number I_π, which provides the measure of the variation of the total of the salaries due to the variations of the individual hourly salaries after we have eliminated the influence of the variations of the number of working days per year as well as the variations of the number of working hours per day.

The procedure is not substantially different from that of the preceding case. We shall deal with the π_i as we have dealt in the preceding case with the p_i, giving to the values of the π_i the corresponding weights $q_i\mu_i$, instead of the weights q_i given to the p_i.

Moreover, if we indicate by $v_{1i}=q_{1i}\mu_{1i}$, $v_{2i}=q_{2i}\mu_{2i}$, the number of

hours worked by the employee i in years 1 and 2, respectively, the problem is reduced to one of finding the index number of the variations of the π_i by eliminating the influence of the variations of the v_i, that is, it is reduced to the elimination of only one group of factors. For some deductions which will follow it is, however, preferable to distinguish the two factors represented by the q_i and the μ_i.

The following expression,

$$I_\pi = \sqrt{\frac{\sum \pi_{2i}q_{1i}\mu_{1i}}{\sum \pi_{1i}q_{1i}\mu_{1i}} \frac{\sum \pi_{2i}q_{2i}\mu_{2i}}{\sum \pi_{1i}q_{2i}\mu_{2i}}}, \tag{2}$$

provides then the index number of the variation of the total annual salaries due to the variations of individual salaries per hour; and the expression,

$$I_{q\mu} = \sqrt{\frac{\sum q_{2i}\mu_{2i}\pi_{1i}}{\sum q_{1i}\mu_{1i}\pi_{1i}} \frac{\sum q_{2i}\mu_{2i}\pi_{2i}}{\sum q_{1i}\mu_{1i}\pi_{2i}}},$$

gives the reciprocal index number of the variation of the total annual salaries due to the variations of the number of working hours per day and of the number of working days per year.

The formulas (1) and (2) may also be written in the forms

$$I_p = \sqrt{\frac{\sum \pi_{2i}q_{2i}\mu_{2i}}{\sum \pi_{1i}q_{1i}\mu_{1i}} \left[\frac{\sum \pi_{2i}q_{1i}\mu_{2i}}{\sum \pi_{1i}q_{2i}\mu_{1i}}\right]}, \tag{1 bis}$$

$$I_\pi = \sqrt{\frac{\sum \pi_{2i}q_{2i}\mu_{2i}}{\sum \pi_{1i}q_{1i}\mu_{1i}} \left[\frac{\sum \pi_{2i}q_{1i}\mu_{1i}}{\sum \pi_{1i}q_{2i}\mu_{2i}}\right]}. \tag{2 bis}$$

The difference between them depends upon the terms put in square brackets. If π_{1i} is constant for all the i, so that we may put $\pi_{1i} = k_1$, and similarly $\pi_{2i} = k_2$ is constant for all the i, formula (2) becomes

$$I_\pi = \frac{k_2}{k_1}, \tag{3}$$

and (1 bis) reduces to

$$I_p = \frac{k_2}{k_1} U, \tag{4}$$

where

$$U = \sqrt{\frac{\sum \mu_{2i}q_{1i}}{\sum \mu_{1i}q_{1i}} \frac{\sum \mu_{2i}q_{2i}}{\sum \mu_{1i}q_{2i}}}, \tag{5}$$

represents the index number, according to the ideal formula, of the variations of the number of working hours per year depending upon

the variations of the number of working hours per day, after we have eliminated the influence of the variations of the number of working days per year.

The above considerations have no practical application for the index numbers of salaries, but are useful in treating some questions relating to the index numbers of prices.

If p_{1i}, p_{2i} represent the prices, q_{1i}, q_{2i} the quantities, and μ_{1i}, μ_{2i} the economic utilities of a commodity i at moments 1 and 2, respectively, π_{1i}, π_{2i} will represent the prices of an economic unit and be the reciprocals of the economic purchasing power of money at the given moments. Then formula (1) provides the usual index number of prices, which is the index number of the physical purchasing power of money, formula (2) provides the index number of the economic purchasing power of money, and formula (5) the index number of the economic utility of the commodities.

But, because of the equilibrium among marginal satisfactions, π_{1i} and similarly π_{2i} are constant for all the commodities; so that the index number of the economic purchasing power of money will be provided by formula (3) and the index number of the physical purchasing power of money by formula (4). The relation,

$$I_\pi = \frac{I_p}{U},$$

shows:

(1) that the index number of the economic purchasing power of money is equal to the index number of the physical purchasing power of money, divided by the index number of the economic utility of commodities;

(2) that, if the index number of the economic utility of commodities has not changed from the moment 1 to the moment 2, the index number of the economic purchasing power of money is equal to the index number of the physical purchasing power of money.

Indicating by

$$P_1 = \frac{\sum p_{2i}q_{1i}}{\sum p_{1i}q_{1i}},$$

Laspeyres' formula of the index number of prices and by

$$P_2 = \frac{\sum p_{2i}q_{2i}}{\sum p_{1i}q_{2i}},$$

Paasche's formula, their geometric mean is

$$I_p = \sqrt{P_1P_2},$$

This expression shows that the index number of prices according to the ideal formula is intermediate between the index numbers according to Paasche's and Laspeyres' formulas. But we have already seen that the index number of prices gives also the index number of the economic purchasing power of money when the index number of the economic utility of commodities is equal to unity. So we may say that, if the index number of the economic utility of commodities equals unity, the formulas of Laspeyres and Paasche give the limits between which the index number of the economic purchasing power of money is contained.

Moreover, we may remark that $P_1 > P_2$ when, between the variations of p_i and q_i from the moment 1 to the moment 2, there is a negative correlation,[5] and that we must expect a negative correlation between p and q because, when a commodity increases in price, its sales tend to be reduced and, on the other hand, when the available quantity of a commodity increases, its price tends to decrease. Correspondingly we must expect that $P_1 > P_2$.

Then we may conclude that, if the index number of the economic utility of the commodities is equal to unity, the Laspeyres' formula gives the upper limit and the Paasche's formula the lower limit of the index number of the economic purchasing power of money.[6]

I do not see the necessity of any other condition.[7]

[5] See, for the demonstration, my article in *Metron:* "Quelques considérations au sujet de le construction des nombres indices des prix et des questions analogues," Vol. 4, No. 1, 1924, pp. 81–83.

[6] This conclusion had already been given in my article cited above, pp. 139–149, but it seems not to have been clear to all readers. Cf. R. Frisch, "Annual Survey of General Economic Theory: The Problem of Index Numbers," in ECONOMETRICA, January 1936, p. 23, and H. Staehle, "A Development of the Economic Theory of Price Index Numbers," in *The Review of Economic Studies*, Vol. 2, No. 3, June 1935, note 2 on page 3 of the reprint. I should add that all my efforts to explain the matter to Dr. Staehle by correspondence, replying to the questions that he had put to me, proved unsuccessful.

[7] The condition of the constancy of the index number of the economic utility of the commodities does not imply the constancy of the economic utility of each commodity; on the contrary the economic utility of some or even of all the commodities may have changed from moment 1 to moment 2, increasing for some of the commodities and decreasing for others. (Cf. the article in *Metron* cited above, page 141). There is consequently no necessity for making the hypothesis that the tastes of the persons remain unaltered from moment 1 to moment 2. (Cf. the article cited, page 148.)

In view of the subjective character of economic utility, it is impossible to speak of the economic utility of a commodity (and consequently of the index number of the economic purchasing power of money) for a group of persons, except on the hypothesis—manifestly contrary to reality—that each commodity, having a certain price, has also the same economic utility for all persons in the group, or at least the ratio between the economic utilities of the various com-

TYPE C

We shall indicate by:

p_{1i}, the antero-posterior diameter of the thorax of the individual i at age 1;

q_{1i}, the transverse diameter of the thorax of the same individual i at age 1;

r_{1i}, the sternal height of the same individual i at age 1;

p_{2i}, the antero-posterior diameter of the thorax of the same individual i at age 2;

q_{2i}, the transverse diameter of the thorax of the individual i at age 2;

r_{2i}, the sternal height of the individual i at age 2.

Consequently, the product $p_{1i}q_{1i}r_{1i}$ indicates the volume of the thorax of the individual i at age 1 and the product $p_{2i}q_{2i}r_{2i}$, his thorax-volume at age 2. The ratio

$$I = \frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}},$$

gives the index of the increase (or decrease) of the volume of the thorax for the population considered from age 1 to age 2.

The problem is now one of finding what part of this variation may be attributed to the variations of the antero-posterior diameter, what part to the variations of the transverse diameter, and what part to the variations of the sternal height.

We may consider in this connection that the value of I may be expressed in the following three different ways

$$(1)\qquad I = \left(\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{1i}r_{2i}}\,\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{2i}r_{1i}}\,\frac{\sum p_{1i}q_{1i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}}\right)^{1/2},$$

$$(2)\qquad I = \left(\frac{\sum p_{1i}q_{2i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{1i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{2i}r_{2i}}\,\frac{\sum p_{2i}q_{1i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\right)^{1/2},$$

$$(3)\qquad I = \left(\frac{\sum p_{2i}q_{1i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{2i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{1i}r_{2i}}\,\frac{\sum p_{1i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\right)^{1/2}.$$

The first factor of the formula (1) may be expanded as follows

$$\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}} = \left(\frac{\sum p_{2i}q_{1i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{1i}q_{2i}r_{1i}}\,\frac{\sum p_{1i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\,\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{2i}q_{1i}r_{1i}}\right)^{1/2}$$

---

modities is constant for all persons of the group. The index number of the economic purchasing power of money may, however, in any case be referred to a single person and it has been suggested by various authorities that it may be properly referred to an *economic average man*, as a representative of the group. (Cf. the article cited, page 129.)

and in a similar way may be expanded the second factor of formula (1) as well as the two first factors of formulas (2) and (3).

If, after such expansions, we calculate the geometrical mean of formulas (1), (2), and (3), we obtain the following formula

$$I = I_p I_q I_r,$$

wherein

$$(4)\quad I_p=\left[\left(\frac{\sum p_{2i}q_{1i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\right)^2\left(\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{2i}r_{2i}}\right)^2\frac{\sum p_{2i}q_{1i}r_{2i}}{\sum p_{1i}q_{1i}r_{2i}}\;\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{1i}q_{2i}r_{1i}}\right]^{1/6},$$

$$(5)\quad I_q=\left[\left(\frac{\sum p_{1i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}}\right)^2\left(\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{1i}r_{2i}}\right)^2\frac{\sum p_{1i}q_{2i}r_{2i}}{\sum p_{1i}q_{1i}r_{2i}}\;\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{2i}q_{1i}r_{1i}}\right]^{1/6},$$

$$(6)\quad I_r=\left[\left(\frac{\sum p_{1i}q_{1i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}}\right)^2\left(\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{2i}r_{1i}}\right)^2\frac{\sum p_{1i}q_{2i}r_{2i}}{\sum p_{1i}q_{2i}r_{1i}}\;\frac{\sum p_{2i}q_{1i}r_{2i}}{\sum p_{2i}q_{1i}r_{1i}}\right]^{1/6},$$

represent respectively the index number of the variation of the thorax-volume due to the variations of the antero-posterior diameter, the index number of the variation of the thorax-volume due to the variations of the transverse diameter, and the index number of the variation of thorax-volume due to the variations of the sternal height.[8]

On the other hand, from formulas (1), (2), and (3), it is easy to derive: the index number I_{qr} of the variation of the volume of the thorax due to the variations of the vertical transverse section, after elimination of the influence of the variations of the antero-posterior diameter; the index number I_{pr} of the variation of the volume of the thorax due to the variations of the vertical antero-posterior section, after elimination of the influence of the variations of the transverse diameter; and the index number I_{pq} of the variation of the volume of the thorax due to the variations of the horizontal section, after elimination of the influence of the variations of the sternal height.

One finds

$$(7)\quad I_{qr} = \left(\frac{\sum p_{1i}q_{2i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}}\;\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{1i}r_{1i}}\right)^{1/2},$$

$$(8)\quad I_{pr} = \left(\frac{\sum p_{2i}q_{1i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}}\;\frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{2i}r_{1i}}\right)^{1/2},$$

[8] The comparison of the expression for I_p given by Wisniewski (cf. note 4), and of the others for I_q and I_r, which may be obtained by substituting q or r, respectively, for p, in the above expressions (4), (5), and (6), leads to the following relations:

$$_wI_p = I_p^{1/2}I^{1/6};\ _wI_q = I_q^{1/2}I^{1/6};\ _wI_r = I_r^{1/2}I^{1/6};$$

where $_wI_p$, $_wI_q$, $_wI_r$ are the values given by Wisniewski and I_p, I_q, I_r are mine.

$$(9) \qquad I_{qp} = \left(\frac{\sum p_{2i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}} \frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{1i}r_{2i}}\right)^{1/2}.$$

The following formulas give the reciprocal index numbers I_p' for the variation of the volume of the thorax due to the variations of the antero-posterior diameter, after elimination of the influence of the variations of the vertical transverse section; I_q', for the variation of the volume of the thorax due to the variations of the transverse diameter, after elimination of the influence of the variations of the vertical antero-posterior section; and I_r' for the variation of the volume of the thorax due to the variations of the sternal height, after elimination of the influence of the variations of the horizontal section.

$$(10) \qquad I_p' = \left(\frac{\sum p_{2i}q_{1i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}} \frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{1i}q_{2i}r_{2i}}\right)^{1/2},$$

$$(11) \qquad I_q' = \left(\frac{\sum p_{1i}q_{2i}r_{1i}}{\sum p_{1i}q_{1i}r_{1i}} \frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{1i}r_{2i}}\right)^{1/2},$$

$$(12) \qquad I_r' = \left(\frac{\sum p_{1i}q_{1i}r_{2i}}{\sum p_{1i}q_{1i}r_{1i}} \frac{\sum p_{2i}q_{2i}r_{2i}}{\sum p_{2i}q_{2i}r_{1i}}\right)^{1/2}.$$

These formulas (10), (11), (12) for the index numbers I_p', I_q', I_r' are not identical with formulas (4), (5), (6) for the index numbers I_p, I_q, I_r obtained after we have eliminated the variations of the other two components separately, rather than the variations of their products. As a matter of fact, in formulas (4), (5), (6), the two first factors correspond to the two factors of formulas (10), (11), (12), but each of formulas (4), (5), (6), has two other factors, in which the variations of the component considered are weighted in a different manner from that adopted in formulas (10), (11), (12).

In the examples of the three types of interrelations given above, we have always considered, for simplicity, series of two terms (populations 1 and 2 in type A; years 1 and 2 in type B; ages 1 and 2 in type C). The methods proposed in this article are, however, applicable to series of any number of terms whatsoever. In the case of series of more than two terms, instead of the ideal formula for the index numbers, the formula for the "circular index number" should be applied.[9]

CORRADO GINI

Harvard University, July, 1936
Cambridge, Mass.

[9] For this formula compare the articles: "Quelques considérations," etc., *loc. cit.*, p. 110, and "On the circular test of index numbers," *Metron*, Vol. 9, No. 2, 1931, p. 10.

MR. KEYNES AND TRADITIONAL THEORY

By R. F. Harrod

In this paper I do not propose to ask or answer the question, has Mr. Keynes succeeded in establishing the propositions which he claims to have established? nor again, what kind of evidence is required to establish or to refute those propositions? I shall confine myself to a narrower question, namely, what are the propositions which Mr. Keynes claims to have established? And in order to restrict my subject matter still further, I propose to confine myself to those propositions, which he claims to have established, that are in conflict with the theory of value in the form in which it has hitherto been commonly accepted by most economists. In other words, my question is what modifications in the generally recognized theory of value would acceptance of the propositions that Mr. Keynes claims to have established entail?

In order to clarify the issues involved it may be well to divide commonly accepted theory into the general theory and its specialised branches. The general theory consists primarily of a number of functional equations expressing individual preference schedules and a number of identities, such as that supply must be equal to demand, and the elucidation of such questions as whether there are as many equations as there are unknowns and whether the solutions are single or multiple. The result of these enquiries should make it clear whether the equilibrium of the system as a whole is stable or unstable or undetermined, whether there are alternative positions of equilibrium, etc. There may be some clues as to the general form of some of the functional equations, provided by such principles as the Law of Diminishing Utility, to use old-fashioned terminology, which may make it possible to predict the direction of changes in the values of the various unknowns due to a given change in one of them. More precise prediction can only be achieved if or when it becomes possible, as a result of the labors of such investigators as Dr. Schultz, to write down the actual terms of the functional equations. Within the corpus of this general theory may be included the formulation of the market conditions that are required for the realisation of some kind of maximum. Thus if one individual A is indifferent whether he produces commodity X or commodity Y for a certain consideration, and another individual B prefers X to Y, the maximum is not realised if the market so operates that A normally produces Y and not X for B. On this condition the maxim of Free Trade is fairly securely founded, the more general maxim of laissez faire much less securely so.

In contrast with the theory of value in this very general form, may be set the special theories formulated to deal with specific problems

such as interest, profit, joint production, discriminating monopoly, etc. The normal method used in dealing with these departmental studies is to assume that certain terms, which appear as variables in the general system of equations, may be treated as constants for the special purpose in hand. For instance in studying the behaviour of duopolistic producers of a given commodity, it may be assumed that the duopolists can obtain the services of factors of production at rates the determination of which in the market will not be appreciably affected by the duopolists' behaviour. Such methods constitute short cuts to the unravelling of particular problems and they are often perfectly legitimate. In the minds of most economists, other than those who stand, so to speak, at the philosophical end of the economic array, the conclusions reached by these short-cut methods constitute the main findings of economic theory.

I may say at once that in my opinion Mr. Keynes' conclusions need not be deemed to make a vast difference to the general theory, but that they do make a vast difference to a number of short-cut conclusions of leading importance. Thus to those whom I may perhaps call without offence the ordinary working economists they ought, if accepted, to appear to constitute quite a revolution. Whether they entail a substantial modification of the more general theory depends on how that is stated. I need hardly observe that there is no authorised version. Those whose main interest is in the general theory, may, if they have laid their foundations well and carefully, be able to look down with a smile of indifference on the fulminations of Mr. Keynes. Pavillioned upon their Olympian fastness, they are not likely to show much irritation.

It is convenient to take Mr. Keynes' theory of interest as the starting point of this exposition. In the commonly accepted short-cut theory there are two unknowns and two equations. The two unknowns are the volume of saving (=the volume of investment) and the rate of interest. Of the new-fangled view, sponsored by some out-of-the-way definitions in Mr. Keynes' *Treatise on Money*, that the volume of saving may be unequal to the volume of investment, it is not necessary to say anything, since it has played no part in the standard short-cut formulations of interest theory (although it has figured in recent writings concerned with practical monetary problems). The commonly accepted interest theory from the time of the early classical writers onward entails that saving is always and necessarily equal to investment.

The two equations in the traditional theory of interest correspond to the demand and supply schedules relating to a particular commodity. First there is the demand equation:

$$y=f(x),$$

y, the marginal productivity of capital, depending on x, the amount of capital invested per unit of time. So much capital will be invested that its marginal productivity is equal to the rate of interest; that is,

$$y=y',$$

where y' is the rate of interest. Since both the traditional theory and Mr. Keynes hold that investment is undertaken up to the point at which the marginal productivity of capital is equal to the rate of interest, y' may be suppressed, and y made to stand for the rate of interest which is equal to the marginal productivity of capital.

Then there is the supply equation:

$$x=\phi(y);$$

x, the amount which individuals choose to save, which is equal to the amount of investment, depends on the rate of interest. Thus there are two unknowns, the rate of interest and the volume of saving, and sufficient equations to determine them. It is not necessary for the present purpose to consider controversies concerning the forms of these equations, such as whether a rise in the rate of interest tends to cause people to save more or less.

This treatment of interest and saving is analogous to that of the price of a particular article and the amount of it produced. The treatment depends on the short-cut assumption of *ceteris paribus*. This is often legitimate in the case of particular commodities, although it is recognised that in certain cases it is idle not to bring in certain other variables, for instance the prices of close substitutes. Among the "other things" which are supposed to be "equal" is the level of income in the community under discussion. In many cases it may be true that when we are trying to determine how much of a particular commodity a producer is likely to produce, his decision to produce a little more or a little less will not have a sufficiently large effect on the total income of the community to react on the market for his goods in such a way as to make an appreciable difference to him. This particular short-cut is in that case justified. I suggest that the most important single point in Mr. Keynes' analysis is the view that it is illegitimate to assume that the level of income in the community is independent of the amount of investment decided upon. No results achieved by the short-cut of such an assumption can be of any value.

How does Mr. Keynes' analysis proceed? His first equation is substantially the same as that of the traditional analysis

$$y=f(x).$$

The marginal productivity of capital is a function of the amount of investment undertaken. The marginal productivity of capital appears in Mr. Keynes' book under the title of marginal efficiency. It does not

appear that there is a difference of principle here. It is true that Mr. Keynes makes an exhaustive and interesting analysis of this marginal efficiency and demonstrates that its value depends on entrepreneurial expectations. The stress which he lays on expectations is sound, and constitutes a great improvement in the definition of marginal productivity. This improvement, however, might be incorporated in traditional theory without entailing important modifications in its other parts.

When we come to the second equation the level of income must be introduced as an unknown term, giving

$$x = \phi(y, i),$$

where i is the level of income. The amount of saving depends not only on the rate of interest, but on the level of income in the community.

It might be thought that to introduce the level of income as an unknown at this point is tantamount to abandoning all attempt to have a departmental theory of the volume of saving, since the level of total income appears in all the equations of the general theory and it is impossible to determine its value without taking all factors into account. This would mean that we should have to leave the ordinary working economist without any departmental theory of saving and interest which he could grasp, and to let him flounder in the maze of $n \times r \times s$, etc., equations governing the whole system. Mr. Keynes has, however, come to the rescue and carved out a new short-cut of his own. In his view the value of the unknown level of income can be determined in a legitimate and satisfactory manner by the departmental equations relating to saving and interest only. To the legitimacy of this assumption it will be necessary to return presently.

Meanwhile, since there are three unknowns and but two equations in the savings/interest complex, another equation is needed. Before proceeding to that, it may be well to recur to the second equation,

$$x = \phi(y, i).$$

This may be transposed into the form

$$i = \psi(x, y).$$

The level of income depends on the amount of investment (=that of saving) and the rate of interest. In this form the second equation shows itself as the doctrine of the multiplier. The multiplier is the reciprocal of the fraction expressing the proportion of any given income, which, at a given rate of interest, people consume. If the value of the multiplier is known for any given rate of interest and level of income, the actual level of income can be deduced directly from the volume of investment. Those to whom the doctrine of multiplier seems an alien morsel in the corpus of economic doctrine should remember that it is

merely a disguised form of the ordinary supply schedule of free capital, but with the level of income treated as a variable.

In discussing this doctrine, for the sake of a still shorter cut, Mr. Keynes is inclined to let the rate of interest drop out of sight. Thus the equation becomes

$$i=\psi(x);$$

the level of income depends on the volume of investment. The justification for this procedure is that whereas the relation of the level of income to the amount of investment is in the broadest sense known—it may be assumed that people save a larger absolute amount from a larger income—the relation of the amount which people choose to save to the rate of interest is a matter of controversy. Moreover in Mr. Keynes' view the level of income has a more important effect on the amount which people choose to save than the rate of interest. However, there is no need to pick a quarrel here. The rate of interest may be brought back into this part of the picture without affecting the main argument. The propensity to consume may be regarded as depending on the rate of interest, although for the sake of brevity and clarity mention of this need not be insisted on at every point in an exposition of the doctrine of the multiplier.

What of the third equation? We have

$$y=\chi(m),$$

where m is the quantity of money, a known term, depending on banking policy. This is the liquidity preference schedule. Probably i, the level of income, ought to be inserted in this equation, thus:

$$y=\chi(m, i),$$

since the amount of money required for active circulation by consumers and traders depends on the level of income. Ought not the price level to come in also? That may be taken to be subsumed under i, the level of income, in a manner that I shall presently explain. The residue of money, not required for active circulation, is available for ordinary people who are discouraged by their brokers from immediate investment, and, more important, for firms, who may want cash for capital extensions or similar purposes within six months or a year or two, and are unwilling to hold their reserves in the form of securities to which some risk of depreciation within the prescribed period is attached. Since the amount of money available for liquid reserves is strictly limited and cannot be increased by the mere desire on the part of firms to hold more money than that, the prospective yield of less liquid reserves must be sufficient to confine those who insist on a money reserve to the amount of money available for that purpose. The less the amount of money available the higher the rate of interest will have to be, both because the high rate is a *quid pro quo* against the risk of

depreciation of the capital and also because the higher the present rate the less probability is there of depreciation within the prescribed period.

It is not necessary to give a final pronouncement on the significance of the liquidity preference equation. It appears that even if some modification is required in this third equation, which determines the rate of interest, a type of analysis similar in its general structure to that of Mr. Keynes may be maintained.

We now have three equations to determine the value of the three unknowns, level of income, volume of saving (=volume of investment), and rate of interest (=marginal productivity of capital).

For the working economist these results may be set out in still briefer shorthand as follows. The amount of investment (=amount of saving) depends on the marginal productivity of capital and the rate of interest; the level of income is connected with the amount of investment by the multiplier, i.e., by the propensity to consume; and the rate of interest depends on the desire for liquid reserves and the amount of spare cash in the community available to satisfy that desire. The amount of this spare cash depends on the policy of the banks in determining the quantity of their I.O.U.'s that are outstanding and on the level of income (the higher this, the more money will be taken away into active circulation).

Thus if the schedules expressing the marginal productivity of capital, the propensity to consume, and the liquidity preference are known and the total quantity of money in the system is known also, the amount of investment, the level of income and the rate of interest may readily be determined.

The next topic for consideration is the legitimacy of the assumption that the level of income may be regarded as determined by the complex of considerations expressed in the savings/interest equations, rather than by the whole system of equations. In general the level of activity is traditionally conceived as depending on the preference schedules of the various factors expressing their willingness to do various amounts of work in return for income, and on the schedules expressing the relation between the amount of work done and the income accruing from it (Laws of Returns). In considering the former schedules we have to take into account all the factors of production. Now in Mr. Keynes' system the supply of capital has already been dealt with by the savings/interest equations. For the supply of risk-bearing, we may provisionally content ourselves with the elegant device which he provides in his footnote to p. 24. He writes, "by his (the entrepreneur's) expectation of proceeds I mean, therefore, that expectation of proceeds, which, if it were held with certainty, would lead to the same

behaviour as does the bundle of vague and more various possibilities which actually makes up his state of expectation when he reaches his decision." Thus considerations affecting the supply of risk-bearing are subsumed in the equations which determine the volume of investment.

There remain the factors other than those covered by the category of investment. Of these we are only concerned with those the supply of which can be varied. Thus we are left with those which may roughly be designated prime factors. What is the nature of their supply-schedule? What is the form of their preference for income in relation to the work required to obtain it?

In this field Mr. Keynes' argument is vitally dependent on his observation of real conditions. The work/income preference schedule exerts its power upon the economic system through the terms on which the prime factors are willing to sell their services. The contracts or bargains of the entrepreneurs with prime factors are normally fixed in money, with no proviso regarding the general level of prices. In the exceptional cases in which there is such a proviso, it is none the less usually the case that a rise in prices involves *some* fall in real rewards to prime factors and conversely. It is true that in a time of rising prices the factors may press for a rise in rewards, but, even if they achieve this, there is still no proviso to safeguard them against a further rise of prices, and prices may, for all the new bargains lay down, and indeed are very often in fact observed to, run on ahead of rewards. Conversely in a time of falling prices. This gives the supply schedules of the prime factors a very special kind of indeterminacy which undermines their power to determine the general level of activity. Mr. Keynes discusses this matter in Book I and its importance in his logical edifice justifies him in giving it pride of place.

Consider next the second set of schedules determining the general level of activity, namely those expressing the relations between the amount of work done and the income accruing from it (Laws of Returns). Since the bargains with prime factors are expressed in money, the returns due to their employment should be expressed in money also. But the money value of these returns depends on the level of prices. The general price level might be regarded as determined by the Quantity Theory of Money; Mr. Keynes does not so regard it for reasons which will be explained below. On the contrary he regards the general price level as completely malleable and determined by the equations in the general field without reference to the quantity of money.

The consequence of the conclusions yielded by the interest/savings equations, if these are accepted, is, that the level of income and activity is determined. Now suppose the entrepreneurs decide to produce more than the amount so determined. Owing to a deficient propensity to

consume, they will find deficient purchasing power, and either accumulate stocks or sell at a loss. If they do the former the accumulation of stocks will constitute an additional (involuntary) investment on the part of the community, which when added to the intended investment, makes the total investment of the community such as to be consistent, in accordance with the interest/savings equations, with the higher level of activity which entrepreneurs are choosing to indulge in. But such a position is unstable. So long as stocks are accumulating, they will reduce activity and continue to do so, until it reaches the point indicated by the interest/savings equations. If on the other hand they sell at a loss, they will be dis-saving; the propensity to consume will be temporarily raised, so that the higher level of activity which they are choosing to indulge in becomes consistent with that required by the interest/savings equations. But again the position is unstable. The marginal propensity to consume will not be permanently sustained at an abnormally high figure by these means. To avoid losses, entrepreneurs will restrict and continue to do so, until activity and income are reduced to a level which satisfies the interest/savings equations, with the marginal propensity to consume normal for that level of income. Converse arguments would apply in the case of entrepreneurs deciding to produce too little.

Now if the level of activity so determined is indeed the equilibrium level of activity, the price level must be appropriate to it. Let us suppose that the price of each commodity is determined by the marginal money cost of production, in the crude way that a tiro might describe erroneously supposing himself to be explaining the true classical theory of cost of production. If the law of diminishing returns prevailed on balance, as Mr. Keynes supposes that it does anyhow in the short period, the general price level would be expected to rise with increases of output and to fall with decreases. To make the matter still more crude and common, suppose prices to vary not merely in proportion to changes in the number of units of factors required per unit of output, as output varies, but also in proportion to changes in rates of reward to the factors. In this case we should find, as output rose and diminishing returns came into play, that the rise of prices would just sufficiently exceed the rise of wages, etc., if any, to cover the increased real marginal cost of production per unit. Factors might press for a rise of rewards, but though they might gain on balance in *some* trades, they would always be beaten by the price level in the system as a whole.

Now this is precisely what Mr. Keynes supposes actually to happen. It is, however, "subject to the qualification that the equality (between marginal cost and price) may be disturbed, in accordance with certain principles, if competition and markets are imperfect" (p. 5). The ob-

jections to this view which upholders of the Quantity Theory of Money might raise must be considered. But first observe its relation to the determination of the level of activity.

Take a period within which prime factor bargains do not change. The supply of each of these in money terms may then be represented by a horizontal straight line. But if prices vary in proportion to costs (cost variations including allowance for overtime rates, the employment of less efficient labour, etc.), then the money value of the marginal net product of each factor must be represented by a co-incident horizontal straight line. Therefore on these conditions the two sets of schedules leave the level of output entirely indeterminate. If the matter is expressed in real terms both sets of schedules are downward sloping to the right; they are still co-incident. If money rewards to factors are raised or lowered in response to changes in the level of employment *and* prices are adjusted accordingly, the same result ensues. Thus this complex of equations does *not* determine the level of activity; therefore it leaves that level free to be determined by the savings/interest complex. Q.E.D.

Thus the crux of the matter seems to have shifted to the Quantity Theory of Money. The essence of the difference between the traditional theory and Mr. Keynes' theory can be put thus: In the traditional theory the supply and demand schedules of all the factors stand on the same footing; the level of activity is an unknown, but the price level is determined by the monetary equation. This determination of the price level enables the level of activity to be determined by the factors' money supply schedules, and by their marginal productivity schedules. In Mr. Keynes' theory the level of activity is determined by the equations governing the savings/interest complex. In the general field, in which we are now only concerned with the demand and supply of prime factors, the level of activity is conceived as determined *ab extra*. It is a known quantity. But the price level is conceived to be completely malleable. If it were not the system in the general field would be overdetermined. Thus the monetary equation is shorn of its former powers. The level of activity being a known quantity the price level is determined by the money cost of production, with suitable modifications for imperfect competition.

What right has Mr. Keynes to gut the monetary equation in this way? Has, then, the banking policy no power to influence the situation? Yes, certainly it has. The fact is that the power residing in the monetary equation has already been used up in Mr. Keynes' system in the liquidity preference equation and it cannot therefore exert any direct influence in the general field. To make it do so would be to use its determining influence twice over. In fact in Mr. Keynes' system all the old pieces reappear, but they appear in different places.

Explanation is necessary. It will be remembered that according to the liquidity preference equation, the rate of interest is determined by the desire of people for liquid reserves and the quantity of money available for that purpose. The quantity of money available for that purpose is equal to the total quantity of money in existence less that required for active trade.[1] Now if the quantity required for active trade were perfectly indeterminate, as it must be by the Quantity Theory—for according to that the price level depends on the quantity of money available for active trade, and therefore it is unknown what quantity of money any given amount of active trade will absorb—the residue would be indeterminate also. But if the m in

$$y = \chi(m)$$

is indeterminate, there are too many unknowns in the interest/savings set of equations. Thus it is necessary to the validity of Mr. Keynes' solution of the problem of investment and interest that the amount of money available for liquid reserves should be determinate, and that involves that the price level should be determined otherwise than by the monetary equation. And so, in Mr. Keynes' system it is.

The matter may be put thus: The savings/interest equations suffice to determine the level of activity, subject to the proviso that the quantity of money which appears in the liquidity preference equation is a known quantity; and this will be known if the price level and therefore the amount absorbed in active trade is known. The equations in the general field suffice to determine the price level, subject to the proviso that the level of activity is known. Thus there is after all mutual dependency. The level of activity will be such that so much money is absorbed in active trade that the amount left over enables interest to stand at a rate consistent with that level of activity.

The mutual interdependency of the whole system remains, but the short-cuts indispensable to thinking about particular problems, as Mr. Keynes has carved them out, remain also.

The amount of investment depends on the marginal productivity of capital and the rate of interest. The level of income and activity is related to the amount of investment by the multiplier, that is by the marginal propensity to consume, the price level is related to the level of activity by the marginal money cost of production (which depends

[1] In his liquidity preference equation Mr. Keynes includes the demand for money for whatever purpose, and the quantity of money that appears in it is the total quantity of money in the community. It has appeared simpler in this part of the exposition to divide this total into two parts, the amount required for active circulation and the residue, to define the quantity of money which appears in the liquidity preference equation as that residue, and the demand which the equation expresses as the demand for purposes other than those of active circulation. This re-definition of terms is merely an expository device and does not imply any departure from Mr. Keynes' essential doctrines.

on the amount of activity undertaken), the amount of money absorbed in active trade depends on the volume of trade and the price level, the amount of money available for liquid reserves is equal to the total amount of money in the system less that required for active trade, and the rate of interest depends on the amount of money available for liquid reserves and the liquidity preference schedule.

It may be well to do some exercises. Suppose the banks to increase the total amount of money available by open market operations. The increment may eventually be divided between active circulation and liquid reserves. An increase of money available for liquid reserves will tend to reduce the rate of interest; and so to increase investment. This will increase the level of income through the multiplier in accordance with the marginal propensity to consume. If the fall in the rate of interest increases the marginal propensity to consume, the increase of income will be *pro tanto* greater, but it is not certain that it does so. The increase of income involves an increase of turnover, and of prices in accordance with the law of diminishing returns. This involves an increased use of money in active circulation. Thus the fall in the rate of interest will not be so great as it would be if all the new money went into liquid reserves. The money will be divided between the two uses, but there is no reason whatever to suppose that the increments in each use will be in proportion to the amounts of money previously employed there, as is assumed in a Quantity Theory using a compendious index number. The comparative size of the increments will depend on the current elasticity of the liquidity preference schedule and the current elasticity of the marginal productivity of capital schedule (which involves expectations).

Suppose a fall in rewards to prime factors. The price level will drop. Money will be released from active circulation for liquid reserves. This will tend to make the rate of interest fall and to react on the level of investment and activity accordingly. Thus the stimulus to activity is very indirect and its effectiveness depends on the same factors as that provided by an increase in the quantity of money. This is very different from the view that a reduction of rewards will stimulate activity because costs fall while prices are sustained by the quantity of money remaining the same.

It appears to me that the achievement of Mr. Keynes has been to consider certain features of traditional theory which were unsatisfactory, because the problems involved tended to be slurred over, and to reconstruct that theory in a way which resolves the problems. The principal features so considered are (1) the assumption that the level of income could be taken as fixed in the departmental theory of interest and saving, (2) the peculiar nature of the supply schedules of the prime

factors which arises out of their bargains being fixed in money without proviso as to the price level, and (3) the failure of monetary theory to explain how the total stock of money is divided between liquid reserves and active circulation, or, in other words, the unsatisfactory character of the theory of velocity of circulation.

I stated above that the old pieces in the traditional theory reappear, but sometimes in new places. It might at first be thought that the liquidity preference schedule is a new piece, and that therefore either the new system is over-determined or the traditional writers must have been wrong in supposing that their system was determined. But it is not really a new piece. The old theory pre-supposed that income velocity of circulation was somehow determined. But precisely how was something of a mystery. Thus the old theory assumed that there was a piece there but did not state exactly what it was. Mr. Keynes' innovation may thus be regarded as a precise definition of the old piece.

By placing it where he does, he overcomes a difficulty, which has been assuming an alarming prominence in recent economic work. In monetary literature the rate of interest has been treated, and increasingly so, as an influence of vital importance in the monetary situation. But in traditional theory, neither in the general system of equations nor in the departmental theory of interest does it appear that the rate of interest is more intimately connected with the *numeraire* than the price of any other factor of production. This is a striking discrepancy. Mr. Keynes introduces the liquidity preference schedule at a point which makes it a vital link between the general system of equations and monetary theory. His treatment is in harmony with recent literature in that he justifies the special connexion of the price of this particular factor with monetary problems. It is an immense advance on recent literature because it removes the discrepancy between the treatment of interest in the two branches of study.

In my judgement Mr. Keynes has not affected a revolution in fundamental economic theory but a re-adjustment and a shift of emphasis. Yet to affect a re-adjustment in a system, which in its broad outlines, despite differences of terminology, has received the approval of many powerful minds, Marshall, Edgeworth, and Pigou, the Austrian School, the School of Lausanne, Wicksell, Pantaleoni, Taussig, and Clark, to mention but a few, is itself a notable and distinguished achievement. And in the sphere of departmental economics and short-cuts, which are of greatest concern for the ordinary working economist, Mr. Keynes' views constitute a genuine revolution in many fields.

The foregoing account has attempted to expound, not to appraise. The only criticism of Mr. Keynes which I venture to offer is that his system is still static. Note has been taken of the fact that at certain

important points, e.g., in his definition of the marginal efficiency of capital, Mr. Keynes lays great stress on the importance of anticipations in determining the present equilibrium.

But reference to anticipation is not enough to make a theory dynamic. For it is still a static equilibrium which the anticipations along with other circumstances serve to determine; we are still seeking to ascertain what amounts of the various commodities and factors of production will be exchanged or used and what prices will obtain, so long as the conditions, including anticipations, remain the same. But in the dynamic theory, as I envisage it, one of the determinands will be the rate of growth of these amounts. Our question will then be, what rate of growth can continue to obtain, so long as the various surrounding circumstances, including the propensity to save, remain the same?

Saving essentially entails growth, at least in some of the magnitudes under consideration. No theory regarding the equilibrium amount of saving can be valid, which assumes that within the period in which equilibrium is established, other things, such as the level of income, do not grow but remain constant.

I envisage in the future two departments of economic principles. The first, the static theory, will be elaborated on the assumption that there is no growth and no saving. The assumption that people spend the whole of their income will be rigidly maintained. On this basis it will be possible to evaluate the equilibrium set of prices and quantities of the various commodities and factors, excluding saving. In the second department, dynamic theory, growth and saving will be taken into account. Equilibrium theory will be concerned not merely with what size, but also with what rate of growth of certain magnitudes is consistent with the surrounding circumstances. There appears to be no reason why the dynamic principles should not come to be as precisely defined and as rigidly demonstrable as the static principles. The distinguishing feature of the dynamic theory will not be that it takes anticipations into account, for those may affect the static equilibrium also, but that it will embody new terms in its fundamental equations, rate of growth, acceleration, de-celeration, etc. If development proceeds on these lines there will be a close parallel between the statics and dynamics of economics and mechanics.

But to develop this theme further would take me too far from my subject.

R. F. Harrod

Christ Church
Oxford, England

REPORT OF THE FIFTH EUROPEAN MEETING OF THE ECONOMETRIC SOCIETY

By Hans Staehle

The 5th European Meeting of the Econometric Society was held in Namur, Belgium, on September 23–26, 1935, at the Château de Namur.[1]

The papers presented may be classified under four headings: (1) theory of economic oscillations, (2) theories of special markets, (3) theories connected with consumers' equilibrium, and (4) special constructions (not, or but loosely, connected with general theory) explaining particular phenomena.

1. THEORY OF ECONOMIC OSCILLATIONS

J. Tinbergen read a paper on "A Mathematical Theory of Business Cycle Policy."[2] It consisted of three parts: (1) the presentation of a simplified business cycle "mechanism," (2) an analysis of its various "influencing coefficients" (*Beeinflussungskoeffizienten*), with a view to discovering those which might be modified by policy, and (3) an analysis of the conditions which would have to be satisfied in order to achieve the aims set by various types of policy.

The "mechanism" consisted of 18 equations connecting 18 variables; 8 of these equations represented definitions (expressing, e.g., the equality of the profit margin per unit of product and of the difference between price and unit cost), while the other 10 expressed "reactions" of some variables on other variables. Several equations contained lags. All equations were given in terms of a system of notation worked out by a group of Dutch mathematical economists. Most of the equations used in his "mechanism" being based on actual statistical enquiries, Tinbergen was able to give some, though of course only a very rough, idea of the size of the "influencing coefficients" in different countries.

Considerable simplification was reached by assuming each variable to consist of an equilibrium value and a (small) deviation from it, and supposing some equilibrium values to have the value 1. By this device, products of equilibrium values and products of an equilibrium value and a deviation can often be simplified, whereas products of deviations may be neglected. Thus all equations become linear.

The mechanism is built upon a number of simplifying assumptions. Thus, business is thought of as consisting of vertically amalgamated enterprises; the influence of stocks is neglected; changes in cash bal-

[1] The attendance is listed in Econometrica, Vol. 4, July, 1936, p. 284.

[2] See the article, "Quantitative Fragen der Konjunkturpolitik," *Weltwirtschaftliches Archiv*, Nov., 1935, pp. 366–399, which contains, among other things, a summary of the paper read at Namur.

ances and in stocks of consumers' goods in the hands of workers are assumed zero; no difference is made between different kinds of finished products (e.g., short-lived and long-lived), or between different kinds of credit (e.g., short-term and long-term); no distinction is made between interest and profits in computing the profit margin; and so on.

By successive elimination the 18 equations may be reduced into a single one, containing only one variable (for which u_t, the "production of consumers' goods started in time unit t," was chosen), and having the form:

$$A_0 u_t + A_1 u_{t-1} + A_2 u_{t-2} + \cdots + A_{17} u_{t-17} = 0.$$

This difference equation describes a movement which will go on as long as no external changes occur. As soon as an external change (a "shock") intervenes, the values of u have to be revised and the further movement must be calculated by means of new "initial values."

Having thus described his "mechanism," Tinbergen pointed out that some of the "influencing coefficients" on which it depends may in fact be changed by man, e.g., those indicating the dependence of wage rates on the volume of production and prices may be manipulated by trade unions, or those giving the relationship of new investment to profit and interest rates may be modified by a public works policy, or by the manipulation of interest rates. It will consequently be possible to exert such influence on several essential spots of the mechanism as would serve particular aims of business cycle policy.

Various types of policy may be distinguished. Tinbergen mentioned a policy attempting to counteract each external shock by another shock; and especially dealt with a policy—more "structural" in character—which aimed at reducing the influence of shocks generally. The latter policy may be described by saying that it aims at the greatest possible dampening effect for each variable, which amounts to postulating that all roots of the equation

$$A_0 x^t + A_1 x^{t-1} + \cdots A_{17} = 0$$

shall have moduli much smaller than unity in order to reduce the sensitiveness of the system against various types of shocks. For the particular mechanism he had described, he arrived at the conclusion that the often postulated non-rigidity of wage rates would not necessarily secure the best policy from a structural point of view.

Professor Tinbergen's paper gave rise to discussion on several points. Replying to Breit, Tinbergen pointed out that the large influence of interest rates upon investment—usually supposed in theory—has not yet been proved statistically to exist. From his own work it

would seem that profit margins were far more important. Taking profit margins *per unit of product*, i.e., deducting the unit cost for labour and materials from the price, he had found prices to lag behind investments by about 6 months, whereas profit margins precede investment by roughly 6 months. Pig iron production was taken as a good index of *new* investments started. Breit thought that replacement investment also exerts an influence. Marschak pointed out that for Germany a clear influence of money rates upon investments had been shown for agricultural machinery. Breit said that, to study the influence of profit margins on investment, the margin should not be taken per unit of product, but per unit of capital—or even per unit of labour, Tinbergen added. It was admitted, after some discussion, that the most correct computation might be made per unit of capacity. Della Riccia thought it was not so much the profit margin per unit, but the total profit in relation to the capital invested, which determined the variations of production, and that expectations of profits were based on volume of sales in the preceding period. As regards volume of savings, Marschak thought that the amounts saved might well turn out to be independent of the rate of interest. Woytinsky quoted hoarding as an example of saving which takes place at zero interest.

Tinbergen's definition of equilibrium, i.e., the situation obtained by substituting constant values in the equations and solving for them, was criticized by Breit who thought it was a special definition in that it related to a stationary situation (*Beharrungszustand*). This discussion was taken up later by Frisch, who proposed definitions of stationary equilibrium, moving equilibrium, disequilibrium, etc.[3]

Further in connection with Tinbergen's paper, Marschak suggested that erratic external shocks might be classified according to the spot where they originate: the sphere of consumption or the sphere of production. Frisch suggested that another classification also could be made, namely, according to whether an unforeseen modification took place (a) in the list of the variables, (b) in the relations between the variables ("influencing coefficients"), or (c) in the "initial conditions" (values of the variables). A further classification would be according to types of shocks. For example, it made considerable difference whether a discontinuous change took place in the amplitude or in the period of a harmonic movement. The former type of shock had far more serious effects.

Marschak doubted the validity of the distinction between cases (a) and (c). As a definition of a shock, Frisch—following up an idea suggested by Divisia at the Leyden meeting—proposed "an event which

[3] Frisch has since published his observations, "On the Notion of Equilibrium and Disequilibrium," *Review of Economic Studies*, Vol. 3, Feb., 1936, pp. 100–105.

prevents the course forecast by theory from producing itself." Tinbergen supported this definition: a particular event might or might not be an external shock, according as the scope of the theory is more or less narrow.

In the course of the discussion, the following concrete examples of shocks were given: the 10% reduction of wages and prices introduced by the Brüning government in Germany, the devaluation of sterling (Tinbergen); the effects of the British coal strike upon Poland (Wisniewski); bankruptcies (Woytinsky); changes in the Belgian price structure following upon the devaluation of the belga (Dupriez). According to Frisch, the long duration of the most recent depression was not due to a cumulation of negative shocks, but was a manifestation of the 50-year long cycle.

Hans Bolza read a paper on "Rhythm in Economic Life."[4] Assuming two trading bodies in a closed economy, he pointed out that the total of means of payment in circulation at any moment t could be represented as the difference between the integrals, taken from 0 to the moment t, of the value of total sales of the two bodies. Fluctuations might occur as between, say, a salaried employee who was paid by the month, and the remaining sector of the economy. In this case, the cash in the hands of the employee was descreasing continuously from the beginning to the end of the month, whereas the "verbal promise" of his employer to pay for his services was increasing continuously during the same period. This, Bolza thought, would give rise to an oscillatory movement of the means of payment involved (cash+verbal promise) through time. This latter point was criticized by Koopmans who pointed out that, since both the cash and the verbal promise curves were assumed straight lines, no discontinuous curve could result from their addition.

2. THEORIES OF SPECIAL MARKETS

M. Breit dealt with the theory of the capital and the money market,[5] and particularly with the question, whether short-term and long-term credit must be considered distinct commodities, having distinct markets. In the literature this question is generally denied, some maintaining that the distinction is only "*privatwirtschaftlich*" in nature, others saying that discrepancies between short and long rates can only be due to market imperfections. This explanation in terms of frictions, however, supposes that it is indifferent to lenders and borrowers whether they supply, or demand, short-term or long-term credit. But such an

[4] See his book, *Der Rhythmus der Wirtschaft*, Berlin, 1935.

[5] See his article, "Ein Beitrag zur Theorie des Geld- und Kapitalmarktes," *Zeitschrift für Nationalökonomie*, Bd. 6, No. 5, pp. 632–659.

assumption would not be admissible for the following reasons: As regards the supply side, it cannot be indifferent to savers (or banks) whether they invest for short or long periods; as regards demand, it is true that there is no peremptory reason (e.g., a legal prohibition) for entrepreneurs to refrain from using short-term credit for the purchase of durable investment goods, or from using long-term credit as circulating capital. But there are a number of factors which prevent the choice from being indifferent, even though the short and long rates may be equal. Some of these factors are of an institutional kind: (1) Mobility is reduced by the existence of costs arising from transformation of long-term into short-term credits, and *vice versa;* (2) to short-term rates the costs of periodical renewal must be added; (3) there may be restrictions of an administrative nature (rationing). Other factors are more truly economic: (4) There is a lower limit, different for each rate, below which it will not fall, which is determined by so-called "objective" or general risks (e.g., the expectation of a moratorium which applies equally to all enterprises); (5) the individual entrepreneur will not always get credit at the market rates, since individual risk elements will be added to them; in addition, the larger the amounts he demands, the higher the rate he will have to pay.

Breit gave a graphic representation of the interaction of these various factors. He made it obvious, for example, how difficult it is to decide which is "the money rate of interest" whose deviations from the natural rate will determine (cumulative) deviations from equilibrium in the distribution of productive resources as between capital and consumers' goods. He further showed that the relatively low level of short term rates during depression is only apparent. Another point was that in case an entrepreneur was working with borrowed capital, the individual risks (as distinct from the general ones) were counted twice (though not necessarily valued at the same amount), namely both by the entrepreneur and by the lender; whereas, in case the producer worked with his own capital, risk was counted but once, i.e., by himself.

This last point was doubted by Tinbergen; but in the ensuing discussion (in which Frisch, Koopmans, and Marschak took part) agreement was reached that in the case of borrowed capital the risk was likely to be overstated.

Marschak delivered a paper "On Investment." Having attempted to measure the net annual investments in several countries, he had become aware of the necessity to agree on a system of definitions relevant to the subject; he had thus worked out a tentative system, some parts of which he presented to the meeting. In the following, only a few points of this paper can be mentioned by way of example.

Assets were defined as complexes of goods or claims; they may also be negative (debts). The value of an asset C_n at point of time n is made up of the discounted value of its future yields:

$$C_n = \frac{a_1}{1 + r_1} + \frac{a_2}{(1 + r_1)(1 + r_2)} + \cdots ,$$

where $a_1, a_2, \cdots$, are the prospective future yields, and $r_1, r_2, \cdots$, the prospective future rates of interest. Investment was defined as the increment of C from point of time n to point of time $n+1$; the value of C at the latter point of time being

$$C_{n+1} = \frac{a_2}{1 + r_2} + \frac{a_3}{(1 + r_2)(1 + r_3)} + \cdots .$$

Investment (S_n) may therefore be written

$$C_{n+1} - C_n = C_n r_1 - a_1 ,$$

where $C_n r_1 = e_n$ is the net receipts of the asset C_n in the time-segment n. From this,

$$e_n = S_n + a_1 ,$$

meaning that the net receipts (or income) of an individual consist of his consumption and investment. The latter may, of course, also be negative; it should include such assets as goodwill. Care must, however, be taken that investment be measured as the difference between such valuations of an asset as are made in the same moment of time. It would, e.g., be wrong to consider as investment the difference between a valuation of assets made last year, and a valuation made this year. It is convenient, therefore, to introduce into the notation double subscripts: $C_{n,n+1}$ meaning, e.g., the value of asset C in moment n, valued in moment $n+1$, i.e., the first subscript relating to the "point of event," and the second to the "point of valuation." This distinction (originally due to Lindahl and Myrdal) between prospective and retrospective magnitudes makes it possible to give a quantitative definition of dynamic profits, i.e., as the difference $(e_{12} - e_{11})$. It consists of two parts: the invested dynamic profit $(S_{12} - S_{11})$, and the consumed dynamic profit $(a_{12} - a_{11})$. Prospective income cannot contain dynamic profits; they are necessarily "retrospective."

In the further parts of his paper, Marschak went on clearing the ground on a great number of other points which are often obscure in the present use of terms made by most writers. He adopted Hicks's suggestion of allowing for an individual's expecting from an asset, for each future time-segment, not one yield x as the only possible, but

rather a whole frequency distribution of possible yields x', x'', $\cdots$, with probabilities p', p'', $\cdots$ $(\Sigma p = 1)$. This frequency distribution he described by its successive moments μ_1, μ_2, $\cdots$, where

$$\mu_1 = \Sigma px,$$

$$\mu_2 = \Sigma p(x - \mu_1)^2,$$

$$\mu_3 = \Sigma p(x - \mu_1)^3, \text{ etc.}$$

The first moment μ_1 was called "expected lucrativeness," the second "expected uncertainty." Marschak further showed, after an analysis of the investor's "tastes" and "opportunities" (alias "obstacles"), how the conditions of equilibrium on the assets market could be derived. In the course of his exposé, he gave quantitative definitions of "impatience," "audacity," "optimism," and "bullishness."

In the discussion, Staehle thought it was difficult to define, e.g., the "audacity" or "bullishness" of the whole market without assuming the frequency distributions, from which these concepts are measured, to be identical for all individuals, because what was "audacious" for one might be "timid" for another. Marschak said these measurements had to be taken from the frequency distributions established by an outside "reasonable" observer, and, in addition, this impartial distribution was required only for the measurement of "bullishness." In reply to Breit, Marschak specified that cash in hand was also an asset.

F. Zeuthen gave a paper on "Effect and Cost of Advertisement from a Theoretic Aspect."[6] He offered a static, partial-equilibrium analysis of advertisement by a firm which has for its product a sloping demand curve, e.g., a monopolist. He first showed that the shifting of the original demand curve due to advertisement is not only a function of the amount spent on advertising, but also of the kind of advertisement chosen. But what interested the seller was not the shift of the demand curve as a whole, but just the movement of that part of the curve which was near the optimum point, at which marginal revenue just equalled marginal costs. The ensuing analysis, therefore, ran in terms of price directions only[7] which were assumed linear, though allowance might easily be made for curvilinear shiftings of the optimum price-quantity points. Zeuthen then showed, by the use of elaborate diagrams, how once the price direction and the course of the marginal cost

[6] The paper has appeared in *Nordisk Tidskrift for Teknisk Økonomi*, No. 1, Sept., 1935, pp. 62–72 (in English).

[7] The question may, of course, be put how this "price direction," depending on the new optimum point, can be found without knowing a considerable portion of the new demand curve.

of advertisement was known (in addition to the original demand and the marginal cost curves), it was possible to determine how far it was profitable to increase advertisement. The equilibrium point is simply determined by the intersection of the marginal revenue curve with the curve obtained by summing together marginal production cost and marginal advertisement cost.

Fernandez Baños gave a summary of his recent paper on the function of the banks in a dynamic economy.

3. THEORIES CONNECTED WITH THE CONDITIONS OF CONSUMERS' EQUILIBRIUM

J. Wisniewski read a paper on "The Elasticity of Demand with Respect to Income,"[8] dealing with price-elasticity and income-elasticity. In this he studied the case of two goods, A and R, whose utilities are assumed independent; such independence could not, he said, be proved statistically, but when it is assumed to exist, the following relations must hold.

Writing

$$\eta = -\frac{dq}{dp}\cdot\frac{p}{q} = \frac{dw}{dp}\cdot\frac{p}{w} + 1 \quad \text{for price elasticity,}$$

$$e = -\frac{dw}{dx}\cdot\frac{x}{w} \quad \text{for income elasticity,}$$

$$\omega = -\frac{dw}{du}\cdot\frac{u}{w} \quad \text{for elasticity of desire (Pigou's concept),}$$

and

$$\Omega = -\frac{dx}{du}\cdot\frac{u}{x}$$

for the reciprocal of the flexibility of money utility, where q stands for quantity, p for price, w for expenditure (pq), x for income, and u for the marginal utility of money, Wisniewski arrived at the relations

$$e = \frac{\omega}{\Omega}\,;\ \eta = 1 + (\omega - 1)(1 - er) \qquad \left[\text{where } r = \frac{w}{x}\right]$$

and

$$e_A r + e_R(1 - r) = 1.$$

From this, he derived the following conclusions: The smaller the value of r, the more closely η approaches ω; on the other hand, $e=\omega$, if $\Omega=1$; so that for $\eta=e$ it is not sufficient to assume r very small, as Marschak

[8] See his article, "Die Elastizität der Nachfrage in Bezug auf das Einkommen," *Zeitschrift für Nationalökonomie*, Bd. 6, No. 3, pp. 386–393.

has done in his "Elastizität der Nachfrage"; Ω will equal unity only by chance. As regards the income-elasticity for all goods combined, it is necessarily equal to unity; this, however, does not mean that it need equal unity for any single good. But an average between the income-elasticities for all goods at least has a clear meaning. This cannot be said for the price-elasticities, η_A and η_B, which are each obtained on the assumption that the price of the other good is constant.

During the discussion, Marschak said that the above equation $e=\omega/\Omega$ had no general validity, because Ω can only exist on certain assumptions regarding the shape of the indifference surface. (For these special conditions see Marschak's contribution to the discussion of Frisch's paper below.) Frisch said, Wisniewski's theory was valuable in that it permitted of obtaining some idea of one of the elasticities when the other was known.

Frisch delivered a paper on "Recent Developments in the Measurement of Money Flexibility." In order to apply the methods of measurement which he had developed in their original form (of which he first gave a succinct summary), it was necessary to use as a "reference commodity" a good, the desire for which was independent of the quantities of all other goods. The first actual measurements had been made by *assuming* the required independence of desire for the reference commodity. These measurements related to members of a Paris co-operative society in 1918–21 (with sugar as a reference commodity), and to families in eight cities in U. S. A. (reference commodity: food), yielding values of the money flexibility of 1.1 to 3.0 and 0.4 to 0.8 respectively. Recently, comparative measurements had become available through collaboration with Frederick V. Waugh and Maurice H. Belz,[9] relating to the U. S. A. as a whole (using yearly data 1919–1931), with butter and meat as reference commodities. They yielded values, at the mid-point of the observations, of 1.11 ($1\pm.05$) via meat, and 1.40 ($1\pm.10$) via butter. In view of this considerable discrepancy, a generalised consumption surface theory was developed, in order to take account of relations of dependence that might exist between the reference commodity and other goods. The new theory runs as follows: Writing p and q for the prices, x and y for the quantities, and u and v for the marginal utilities of two distinct, but related, commodities; P for the price of living, r for real income, and w for the marginal utility of money; the following fundamental relations are assumed:

$$\frac{u(x, y)}{p}=\frac{v(x, y)}{q}=\frac{w(r)}{P}.$$

[9] See Ragnar Frisch, *Statistical Confluence Analysis by Means of Complete Regression Systems*, Oslo, 1934, pp. 147 ff.

Putting $p/q=\lambda$, and $P/p=\alpha$ (the "relative cheapness of x") we see that the first of the above equations defines y in terms of x and λ, $y=f(x, \lambda)$. Therefore, utilizing the equality between the first and last members, we get

$$w(r) = \alpha \cdot u[x, f(x, \lambda)]$$

or

$$w(r) = \alpha \cdot U(x, \lambda),$$

$U(x, \lambda)$ being the function $u[x, f(x, \lambda)]$. In order to apply the isoquant principle to the consumption surface defined by the last equation, it is necessary to use two points for which not only x, but also λ is constant, i.e., two points must be found for which $x_1=x_2$ and $\lambda_1=\lambda_2$. The latter will be true automatically if λ is constant throughout, e.g., if the two commodities are connected by a rigid relation of substitution. The isoquant method in its original form will therefore be applicable whenever a close proportionality exists between p and q. When the relation between the commodities is of the more general form $F(p, q)=0$, the equation may be written

$$w(r) = \alpha \cdot (Ux,p)$$

where p is the absolute price. In the latter case, the inclusion of p takes care of relations not only with the second commodity, but with any number of other goods; the only assumption being that there exists between p and the prices of each of these other goods some definite functional relationship. By the application of this more general consumption surface theory, the discrepancies between the flexibility measurements were practically eliminated; for instance, the money flexibility in U. S. A. 1919–1931 via meat and butter became respectively 0.95(1 ± .05) and 0.97(1 ± .20).

In a long contribution which really amounted to a paper, Marschak drew attention to certain restrictions to which Frisch's consumption surfaces are subject. Frisch's theory, he said, started out by adding one term to the ordinary consumers' equilibrium equations, as follows:

$$\frac{u}{p} = \frac{v}{q} = \cdots = \frac{\phi(r)}{P}$$

where $r = \dfrac{px+qy}{P}$ = real income, and P the "price of living." The usual practice adopted for the computation of P was to use Laspeyres' index number formula. Marschak's following demonstration was to show that, in the case of Laspeyres' formula, the consumption surface could

not exist, whereas when Paasche's index formula was used the surface may exist but is subjected to peculiar restrictions.

Let $f(x, y)$ be a utility surface, and $f_x = u(x) = u$ and $f_y = v(y) = v$ be the marginal utilities of the two independent commodities. Write for the above Frisch's equations:

$$\frac{u}{p} = \frac{v}{q} = \frac{\phi\left(\dfrac{px + qy}{P}\right)}{P}. \tag{1}$$

Calculate P first by Paasche's index formula, i.e.,

$$P = \frac{px + qy}{p_0x + q_0y}$$

choosing such units as to make $p_0 = q_0 = 1$. Then equation (1) transforms into

$$\frac{u}{p} = \frac{v}{q} = \frac{(x + y)\cdot\phi(x + y)}{px + qy};$$

$$\therefore \quad xu + yv = (x + y)\cdot\phi(x + y). \tag{2}$$

The solution of this differential equation is:

$$f(x, y) = h\left(\frac{x}{y}\right) + \int \phi(x + y)d(x + y)$$

where h is an arbitrary function.

If on the other hand, Laspeyres' index formula is used, so that

$$P = \frac{px_0 + qy_0}{p_0x_0 + q_0y_0}$$

no equation independent of p and q—as in (2) above—is obtainable. Therefore ϕ cannot exist in this case.

As to the economic meaning of the solution in the Paasche case, it comes to this: total utility depends both on real income, and on the shares (quantity ratio) of the two commodities, but in such a way that total utility reacts on changes in real income independently of how this income was previously distributed between different commodities; and it reacts on changes in this distribution, independently of the prevailing level of real income. For example, a given change in the quantity ratio of motor cars to bread would affect total utility in the same way (assuming, of course, there were no other commodities), whether the individual was rich or poor. Allen thought that Marschak's restrictions on the utility surface were vaguely connected with those implied

in the "income proportionality"[10] assumption as made by Haberler in his book on index numbers. Marschak, he said, at any rate had shown that one *must use Paasche's* and not the usual Laspeyres' index.

H. Staehle delivered a paper on "A Method of Measuring Variations in the Price of Living, with an Application to Belgium 1891–1932." Starting from Laspeyres' and Paasche's index formulae ($P_0 = \Sigma p_1 q_0 / \Sigma p_0 q_0$, and $P_1 = \Sigma p_1 q_1 / \Sigma p_0 q_1$, respectively, where p_0', p_0'', $\cdots$; p_1', p_1'', $\cdots$ are the prices, and q_0', q_0'', $\cdots$; q_1', q_1'', $\cdots$ are the quantities relating to two situations 0 and 1), Staehle showed that P_0 is higher and P_1 lower than the "true" price of living either if the two incomes that supplied the q_0's and q_1's for the calculation of P_0 and P_1 are equivalent (and this was the case studied by Bortkiewicz, Keynes, and R. G. D. Allen); or if the "income-proportionality" assumption is made, i.e., one supposes that the "true" price of living was the same both on the psychic-income level to which the q_0's belong, and on the other psychic-income level to which the q_1's belong (this was Haberler's case). In these circumstances, the theory of "limits" is either superfluous (since the true price-of-living index could be obtained directly as the ratio of the two equivalent incomes—and the "limits" had only an ornamental value), or it is based on an assumption which, as can be shown by experiment, is in contradiction with facts: the price of living is as a matter of fact different on different income levels, and a theory which assumes these differences away, suffers at least from a lack of generality.

Staehle then passed on to explain that it was possible to avoid both these difficulties by the use of a theory which had been developed by a Russian writer, A. A. Konüs, in 1924. His argument was to show that, while P_0 was always larger than the true price of living (I_{01}) applying to the psychic-income level corresponding (at given tastes) to the q_0's, and P_1 was always smaller than the "true" price-of-living index ($1/I_{10}$) corresponding (at the same tastes) to the q_1's, so that

$$P_0 \geqq I_{01}; \quad \text{and} \quad P_1 \leqq 1/I_{10}.$$

The conclusion of the theory of limits, i.e.,

$$P_0 \geqq I_{01} \geqq P_1, \tag{3}$$

was possible only when one had

$$I_{01} = 1/I_{10}.$$

To have this condition fulfilled, it was necessary, apart from making the above Allen's or Haberler's assumptions, to have one or other of the following cases:

[10] The term "income proportionality" was introduced by Frisch; see his survey in ECONOMETRICA, Vol. 4, Jan., 1936, pp. 1–38.

$$(4) \qquad \Sigma p_0 q_1 = \Sigma p_0 q_0, \text{ or}$$

$$(5) \qquad \Sigma p_1 q_0 = \Sigma p_1 q_1.$$

When P_0 and P_1 were computed by the use of data for the q_0's and the q_1's fulfilling with respect to each other either of conditions (4) and (5), then the conclusion (3) could be drawn: in the case the data fulfilled condition (4), the price-of-living index (to which P_0 and P_1 were the limits) related to the psychic-income level of the q_0's; in the case of condition (5), P_0 and P_1 were limits to the price-of-living index relating to the psychic-income level of the q_1's.[11] This theory of Konüs thus offers the possibility of determining by means of an *objective* criterion the particular money-incomes in two situations, 0 and 1, from which the q_0's and q_1's must be taken in order to arrive at correct limits, provided tastes, environment, etc., could be assumed identical in the two situations.

Staehle then produced an application of this theory to data relating to Belgian workers' families at five different periods of time (1891, 1908, 1921, 1929, and 1932). He dealt with the difficulty arising from the fact that family budget enquiries generally give data on quantities and prices for food items only, and devised the assumptions and technique by which the resulting values of P_0 and P_1 (based on food only) could be used to find the price of living as a whole. He emphasized that his results, which were not obtained on the basis of mechanical considerations, but purported to be limits to *the* price of living on given psychic-income levels, had to fulfill the circular test, and showed that they actually did. Finally he made a comparison of his own price-of-living indices with the official Belgian cost-of-living indices over the periods 1921–29 and 1921–32, showing that the latter underestimated the rise in the cost of living in very considerable proportions. His results also showed considerable differences in the price-of-living indices on different income levels.

The discussion was opened by Frisch who restated the main conclusions of the paper in a rather more general way. He was not convinced of the objection raised by Staehle against Haberler's "income-proportionality." The error involved, he thought, could not be very great, provided the range between the incomes supplying the q_0's and the q_1's was not too wide (in terms of either money scale). He did admit, however, that for larger variations Staehle's objection was important.

Allen suggested that, in order to get a more continuous set of ob-

[11] For a full proof see Staehle's article: "A Development of the Economic Theory of Price Index Numbers," *Review of Economic Studies*, Vol. 2, June, 1935, pp. 163–188.

servations, it might be worth while to interpolate the observed figures for $\Sigma p_0 q_0$, $\Sigma p_1 q_0$, $\Sigma p_0 q_1$, and $\Sigma p_1 q_1$, so that more couples of incomes fulfilling Konüs' conditions (4) or (5) could be found.

Woytinsky thought the method was not of practical value, for even supposing family budget enquiries could be run at sufficiently short intervals, the results would be known too late. And if in the meantime an index of the simple form P_0 was published, it was impossible afterwards to revise these figures, for practical reasons (e.g., sliding scales). Staehle replied that the question was whether or not accurate measurements were wanted; if they were really wanted, then he did not see any possibility of avoiding the use of rather complicated methods.

4. SPECIAL CONSTRUCTIONS

In a paper "Recherches analytiques sur le développement normal des fortunes et des revenus des nations," A. Della Riccia undertook to explain a relationship between national incomes and national wealth (both taken in gold and per capita) he had met with in a recent book of his.[12] Statistics relating to some 35 countries had then led him to the conclusion that, as between these countries, the income per capita (R) could well be represented in terms of wealth per capita (F) by the function $R=a+bF^3$, so that the ratio R/F was increasing with increasing values of F.

Starting from a classification of a nation's resources into those having the nature of (limited) stocks—exhaustible natural resources, like minerals, and existing capital goods—and others having the nature of (perpetual) flows—non-usable or self-perpetuating natural resources and human labour—quantitative definitions were given in terms of these resources of the nation's income and wealth at different points of time (year 0 and year n), and of their yearly increments. In the development of these functions, such coefficients as the proportion of annual income saved, the proportion of capital equipment to be replaced yearly, etc., were taken as constants—though, it was pointed out, had they been assumed variable, the conclusions would not have been modified essentially. Taking into account that every year's production involved a reduction of the remaining total of stocks, Della Riccia arrived at the following conclusions: (1) For the income R to increase in time it is necessary and sufficient that the capital equipment increases every year, i.e., that saving be larger than replacement of capital equipment; (2) that the average annual addition to wealth was decreasing the longer the period over which the average was taken extended into the future, so that dF/dt was a decreasing

[12] *Recherches et opinions économiques et sociales d'un amateur de chiffres*, Bruxelles, 1933, pp. 26–29.

function, tending toward zero; (3) that income was describing in the course of time a logistic function; (4) as a consequence, the ratio R/F was first diminishing, then increasing and finally becoming stable.

In the discussion Woytinsky found the agreement between this theory and the statistics for different countries striking. Frisch thought the logistic curve describing growth of income—one of Della Riccia's results—might have been derived more easily.

Wl. Woytinsky read a paper on "Three Sources of Unemployment and their Measurement." Putting S for the total number of wage-earners in employment, Ch for the number unemployed, V for the volume of production, and T for the output per capita of the persons employed, he gave the following definitions:

$$E = V/T; \quad \text{and} \quad Ch = S - E = S - V/T.$$

From these he derived for an increment of Ch:

$$\Delta Ch = \Delta S + E\frac{\Delta T}{T + \Delta T} - \frac{\Delta V}{T + \Delta T}.$$

In this equation ΔS represents the increment of the active population which, but for changes in T and V, would have added to unemployment; $E\Delta T/T+\Delta T$ stands for the number of wage earners who, assuming V had remained constant, would have been dismissed in consequence of the increased output per head which has risen by ΔT; $\Delta V/T+\Delta T$ represents the number of wage earners who were able to find employment at the new efficiency level, on account of the increase in production ΔV. A given increase in unemployment can thus be resolved into three component parts, which can easily be determined statistically, when the initial values of S and V are taken as 100. If unemployment (Ch) was insignificant at the initial date, it may be supposed equal zero. Then $E=S=100$. Statistical examples were then offered of the decomposition of unemployment, particularly for Great Britain, where unemployment in the years 1870, 1880, etc., up to 1910 was "explained" in comparison to the situation prevailing in 1860. A vast number of applications which were recently published[13] was summarized, and some considerations as to the desirable improvements of various kinds of statistics were added.

During the discussion, Nixon said that E, at least for England, also depended on hours, which fell heavily between 1860 and 1910, to which Woytinsky replied, that was taken care of by the calculation of efficiency. Bowley said the years 1860, 1870, etc. chosen represented

[13] "Three Causes of Unemployment," International Labour Office, *Studies and Reports*, Series C, No. 20, Geneva, 1935.

quite different phases of the trade cycle, and that the difficulty might be overcome by smoothing. He also discussed various details of British statistics which had been touched upon by Woytinsky, and suggested that, in splitting industrial from non-industrial employment, the building trades should be lumped together with industry. The increase in active population would stop before the total population would become stationary. Marschak thought that the second term in Woytinsky's equation did not sufficiently describe the influence of efficiency. He suggested T should be taken as a function of both V and a parameter characteristic of the "technical level."

This report of the Namur meeting would be quite incomplete, if no mention were made of the fine spirit which ruled during the whole session. Everybody was discussing intensely, both during the formal sessions and outside. The attendance of a relatively large number of young newcomers was particularly gratifying. The meeting was altogether a brilliant success, both for the high standard of the papers and discussions, and for the fresh enthusiasm that everybody brought home from it.

Hans Staehle

International Labour Office
Geneva

NOTE ON A NEW EDITION OF THE WORKS OF LÉON WALRAS

MEMBERS of the Econometric Society will welcome the announcement of a plan to publish, as rapidly as available resources and the response of the public permit, a definitive edition of the writings of Léon Walras.

As a first step in this direction, Messrs. Pichon et Durand-Auzias, Paris, have published, at fifty francs per volume, a new edition of the *Etudes d'économie sociale* (last published in 1896) and the *Etudes d'économie politique appliquée* (last published in 1898). The latter contains, in addition to a portrait of Walras and a brief prefatory note by M. Gaston Leduc, a page of errata, drawn up by Walras himself. At least one of these corrections—the substitution of the phrase "l'encaisse désirée" for "la circulation à desservir" (p. 3)—is of importance, since it provides additional confirmation of the validity of the account given in recent years of the evolution of Walras's thought on the subject of the value of money.

The new edition of the *Etudes d'économie sociale* is of particular interest because it includes, in conformity with Walras's express desire, the paper entitled *Souvenirs du Congrès de Lausanne*, which Walras had omitted from the original edition of the *Économie sociale* because of his dissatisfaction with certain of the arguments advanced therein. Some interest attaches to Walras's change of attitude on the points in question (pp. 377 f.); but there can be little doubt that the chief interest of the essay for contemporary readers lies in the light it throws upon the origin of Walras's academic career.

The essay should be read in conjunction with the report of the Congress of Lausanne given by Joseph Garnier in the *Journal des économistes* for October, 1860 (especially pp. 79 and 81 f.). To Garnier, Walras was obviously an impertinent young upstart whose "success" in disturbing the meeting was due "first, to his talent for speaking; secondly, to his boldness; thirdly, to the tone, now dogmatic and now paradoxical, which he gave to his exposition, and to the formulas which he sprinkled through it." For the rest, M. Garnier was prepared to quote with approval the "witty" characterization of Walras's paper, made by an inconsequential German delegate to the Congress, as an "economic charade," as well as the jibe at Walras made by another member of the Congress: "Those who applaud you have not understood you."

The rest of the story is told in Walras's preface to the present edition of the *Souvenirs* (p. 378). One of those who "applauded but did not understand"—a young lawyer and politican by the name of Louis Ruchonnet, who sat in the public gallery and did not even make Walras's acquaintance at the Congress—sought him out in Paris to express his admiration of the one thing he could understand: namely, Walras's

contention that wages should be exempt from taxation. Ten years later, the young lawyer had become head of the Department of Public Education at Lausanne; and it was he who undertook to create a chair in economics at Lausanne if Walras would compete for it.

If this much could have been done for Walras himself by one who "applauded but did not understand," something can be done for his memory by those who continue to applaud and at least pretend to understand. The beginning of the present project has been made possible by the virtually unaided efforts of Walras's daughter, Mlle. Aline Walras, at an almost incredible personal sacrifice. Its continuation—which includes a plan for the publication, under the title *Mélanges d'économie politique et sociale*, of a volume of selected papers by Walras, most of them now difficult of access, and of Walras's own *Abrégé des Éléments d'économie politique pure*, hitherto unpublished—is entirely dependent upon the response made by economists in general, and members of the Econometric Society in particular, to the two volumes now offered to the public.

ARTHUR W. MARGET

University of Minnesota
Minneapolis, Minnesota

THE SUMMATION OF RANDOM CAUSES AS THE SOURCE OF CYCLIC PROCESSES*

By Eugen Slutzky

I. SCOPE OF THE INVESTIGATION

Almost all of the phenomena of economic life, like many other processes, social, meteorological, and others, occur in sequences of rising and falling movements, like waves. Just as waves following each other on the sea do not repeat each other perfectly, so economic cycles never repeat earlier ones exactly either in duration or in amplitude. Nevertheless, in both cases, it is almost always possible to detect, even in the multitude of individual peculiarities of the phenomena, marks of certain approximate uniformities and regularities. The eye of the observer instinctively discovers on waves of a certain order other smaller waves, so that the idea of harmonic analysis, viz., that of the possibility of expressing the irregularities of the form and the spacing of the waves by means of the summation of regular sinusoidal fluctuations, presents itself to the mind almost spontaneously. If the results of the analysis happen sometimes not to be completely satisfactory, the discrepancies usually will be interpreted as casual deviations superposed on the regular waves. If the analyses of the first and of the second halves of a series give considerably divergent results (such as, for example, were found by Schuster while analyzing sunspot periodicity),[1] it is, even then, possible to find the solution without giving up the basic concept. Such a discrepancy may be the result of the interference of certain factors checking the continuous movement of the process and substituting for the former regularity a new one which sometimes may

* Professor Eugen Slutzky's paper of 1927, "The Summation of Random Causes as the Source of Cyclic Processes," *Problems of Economic Conditions*, ed. by The Conjuncture Institute, Moskva (Moscow), Vol. 3, No. 1, 1927, has in a sense become classic in the field of time-series analysis. While it does not give a complete theory of the time shape that is to be expected when a given linear operator is applied to a random (auto-non-correlated) series, it has given us a number of penetrating and suggestive ideas on this question. It has been, and will no doubt continue to be, highly stimulating for further research on this vast and—not least for business-cycle analysis—most important problem. Unfortunately Professor Slutzky's paper so far has been available only in Russian (with a brief English summary). Some years ago Professor Henry Schultz had the original article translated into English by Mr. Eugene Prostov, and suggested that it be published in Econometrica. At the request of the Editor Professor Slutzky has prepared for our Journal a revised English version with which he has incorporated also a number of important results obtained after 1927.—Editor.

[1] Arthur Schuster, "On the Periodicities of Sunspots," *Phil. Trans.*, Series A, Vol. 206, 1906, p. 76.

even happen to be of the same type as the former one. Empirical series are, unfortunately, seldom long enough to enable one definitely to prove or to refute such an hypothesis. Without dwelling on the history of complicated disputes concerning the above-mentioned problem, I will mention only two circumstances as the starting points for the present investigation—one, so to speak, in the field of chance, the other in the field of strict regularity.

One usually takes the analysis of the periodogram of the series as the basis for the discovery of hidden periodicities. Having obtained from the periodogram the values of the squares of the amplitudes of the sinusoids, calculated by the method of least squares for waves of varying length, we ask whether there is a method of determining those waves which do not arise from chance. Schuster apparently has discovered a suitable method;[2] but we must give up his criterion when we remember that among his assumptions is that of independence of the successive observations. As a general rule we find that the terms of an empirical series are not independent but correlated and at times correlated very closely. This circumstance, as is known, may very perceptibly heighten the oscillation of the derived characteristics of the series, and it is quite conceivable that waves satisfying Schuster's criterion would in fact be casual—just simulating the presence of a strict regularity.[3] Thus we are led to our basic problem: is it possible that a definite structure of a connection between random fluctuations could form them into a system of more or less regular waves? Many laws of physics and biology are based on chance, among them such laws as the second law of thermodynamics and Mendel's laws. But heretofore we have known how regularities could be derived from a chaos of disconnected elements because of the very disconnectedness. In our case we wish to consider the rise of regularity from series of chaotically-random elements because of certain connections imposed upon them.

Suppose we are inclined to believe in the reality of the strict periodicity of a business cycle, such, for example, as the eight-year period postulated by Moore.[4] Then we should encounter another difficulty. Wherein lies the source of the regularity? What is the mechanism of

[2] A. Schuster, "On the Investigation of Hidden Periodicities, etc.," *Terrestrial Magnetism*, Vol. 3, 1898.

[3] The further development of Schuster's methods, which we find in his extremely valuable paper, "The Periodogram of the Magnetic Declination as Obtained from the Records of the Greenwich Observatory during the Years 1871–1895," *Trans. of the Cambridge Philos. Soc.*, Vol. 18, 1900, p. 107, seems to overcome this difficulty. Because it is rather unfinished in mathematical respects, however, the influence of this paper seems not to have been comparable to its importance.

[4] H. L. Moore, *Generating Economic Cycles*, New York, 1923.

causality which, decade after decade, reproduces the same sinusoidal wave which rises and falls on the surface of the social ocean with the regularity of day and night. It is natural that even now, as centuries ago, the eyes of the investigators are raised to the celestial luminaries searching in them for an explanation of human affairs. One can dauntlessly admit one's right to make bold hypotheses, but still should not one try to find out other ways?[5] What means of explanation, however, would be left to us if we decided to give up the hypothesis of the superposition of regular waves complicated only by purely random components? The presence of waves of definite orders, the long waves embracing decades, shorter cycles from approximately five to ten years in length, and finally the very short waves, will always remain a fact begging for explanation. The approximate regularity of the periods is sometimes so distinctly apparent that it, also, cannot be passed by without notice. Thus, in short, *the undulatory character of the processes and the approximate regularity of the waves* are the two facts for which we shall try to find a possible source in random causes combining themselves in their common effect.

The method of the work is a combination of induction and deduction. It was possible to investigate by the deductive method only a few aspects of the problem. Generally speaking, the theory of chance waves is almost entirely a matter of the future. For the sake of this future theory one cannot be too lavish with experiments: it is experiment that shows us totally unexpected facts, thus pointing out problems which otherwise would hardly fall within the field of the investigator.[6]

II. COHERENT SERIES OF CONSEQUENCES OF RANDOM CAUSES AND THEIR MODELS

There are two kinds of chance series: (1) those in which the probability of the appearance, in a given place in the series, of a certain value of the variable, depends on previous or subsequent values of the variable, and (2) those in which it does not. In this way we distinguish

[5] A similar viewpoint is found in the remarkable work of G. U. Yule, "Why Do We Sometimes Get Nonsense-Correlations between Time Series?" *Journal of the Royal Statistical Society*, Vol. 89, 1926. This work approaches our theme rather closely.

[6] The following exposition is based on a large amount of calculation. The author expresses special gratitude to his long-time collaborator, E. N. Pomeranzeva-Ilyinskaya and also to O. V. Gordon, N. F. Rein, M. A. Smirnova and E. V. Luneyeva. The calculations were carefully checked, almost all work having been independently performed by two individuals. It is very unlikely that undetected errors are sufficiently significant to affect to any perceptible degree our final conclusions. A few errors, detected in the course of time in Tables I, III, and IX of the original paper, are noted at the end of this paper, and an error in Figure 7, B_4, has been corrected when it was re-drawn.

between *coherent*[7] and *incoherent* (or random) series. The terms of the series of this second kind are not correlated. In series in which there is correlation between terms, one of the most important characteristics is the value of the coefficient of correlation between terms, considered as a function of the distance between the terms correlated. We shall call it the *correlational function* of the corresponding series and shall limit our investigation to those cases in which the distribution of probabilities remains constant. The coefficient of correlation, then, is exclusively determined by the distance between the terms and not by their place in the series. The coefficient of correlation of each member with itself (r_0) will equal unity, and its coefficient of correlation (r_t) with the tth member following will necessarily equal its coefficient (r_{-t}) with the tth member preceding.

Any concrete instance of an experimentally obtained chance series we shall regard as a *model* of empirical processes which are structurally similar to it. As the basis of the present investigation we take three models of purely random series and call them the first, second, and third basic series. These series are based on the results obtained by the People's Commissariate of Finance in drawing the numbers of a government lottery loan. For the first basic series, we used the last digits of the numbers drawn; for the second basic series, we substituted 0 for each even digit and 1 for each odd digit; the third basic series was obtained in the same way as the second, but from another set of numbers drawn.[8]

Let us pass to the coherent series. Their origin may be extremely varied, but it seems probable that an especially prominent role is played in nature by the process of *moving summation* with weights of one kind or another; by this process coherent series are obtained from other coherent series or from incoherent series. For example, let causes $\cdots x_{i-2}, x_{i-1}, x_i, \cdots$ produce the consequences $\cdots y_{i-2}, y_{i-1}, y_i, \cdots$, where the magnitude of each consequence is determined by the influence, not of one, but of a number of the preceding causes, as for instance, the size of a crop is determined, not by one day's rainfall, but by many. If the influence of causes in retrospective order is expressed by the weights $A_0, A_1, A_2, \cdots A_{n-1}$, then we shall have

$$(1)\quad \begin{cases} y_i = A_0x_i + A_1x_{i-1} + \cdots + A_{n-1}x_{i-(n-1)}, \\ y_{i-1} = \qquad\quad A_0x_{i-1} + \cdots + A_{n-2}x_{i-(n-1)} + A_{n-1}x_{i-n}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases}$$

[7] I venture to propose this name because it seems to me that it truly expresses what is intended, namely, the existence of some connection between the elements or parts of a thing (for example, of a series), but not a connection between this thing as a whole and another.

[8] The tables giving these series and seven others derived from them will be found in the original paper (*loc. cit.*, pp. 57–64) and are not repeated here.

Each of two adjacent consequences has one particular cause of its own, and $(n-1)$ causes in common with the other consequence. Because the consequences possess causes in common there appears between them a correlation even though the series of causes are incoherent. When all the weights are equal (*simple* moving summation) the coefficient of correlation expresses the share of the common causes in the total number of independent causes on which the consequences

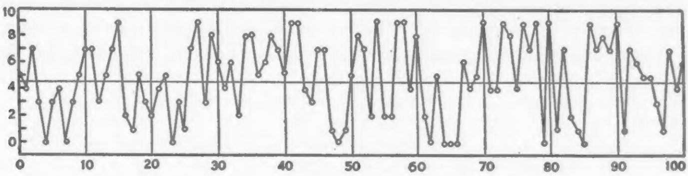


FIGURE 1.—The first 100 terms of the first basic series.

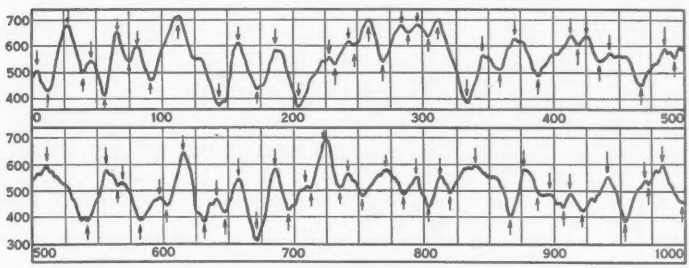


FIGURE 2.—The first 1000 terms of Model II.

depend (as has long been known from the theory of the experiment of Darbishire); then

$$r_0 = 1, r_1 = r_{-1} = \frac{n-1}{n}, r_2 = r_{-2} = \frac{n-2}{n}, \quad \cdots, r_{n-1} = r_{-(n-1)} = \frac{1}{n},$$

further coefficients being equal to zero. By taking a ten-item moving summation of the first basic series, Model I was obtained.[9] A small section of Model I is plotted in Figure 3 with an index[10] of the English

[9] In addition, 5 was added to each sum. This does not change the properties of the series. Neither does it make any difference as to the method of numbering consequences in comparison with the scheme used in formula (1). At the outset of the work, it seemed to be technically more convenient to give the consequence the same number as the earliest cause and not the latest. Thus, for example, for Model I,

$$y_0 = x_0 + x_1 + x_2 + \cdots + x^9 + 5.$$

[10] Dr. Dorothy S. Thomas, "Quarterly Index of British Cycles," in W. L. Thorp, *Business Annals*, New York, 1926, p. 28.

business cycles for 1855–1877 in juxtaposition—an initial graphic demonstration of the possible effects of the summation of unconnected causes.

In turn the consequences become causes. Taking a ten-item moving summation of Model I, we obtained the 1000 numbers of Model II. Performing a two-item moving summation twelve times in succession on the third basic series,[11] we obtained the 1000 numbers of Model IVa. First and second differences of Model IVa give Models IVb and IVc respectively (See Figure 4). Furthermore, the application of scheme (1) to the second basic series gives[12] Model III if the weights used are

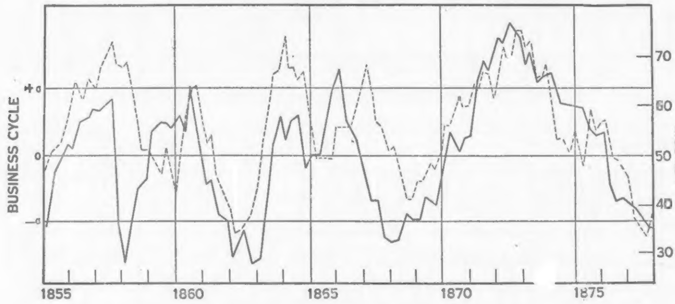


FIGURE 3.————An index of English business cycles from 1855 to 1877; scale on the left side. ———Terms 20 to 145 of Model I; scale on the right side.

10^4 times the ordinates of the Gaussian curve taken at intervals of 0.1σ. Because this model was very smooth it appeared sufficient to use only the 180 even members out of the 360 items (see Figure 11 under the numbers 0, 2, 4, $\cdots$ 358). Model IIIa—the last one—is 10^4

[11] It actually was computed by applying the scheme (1) to the third basic series with the weights 1, 12, 66, 220, 495, 792, 924, 792, $\cdots$, 12, 1, because s-fold simple summation of two items is equivalent, as can be shown easily, to direct summation with the weights $C_0^s, C_1^s, C_2^s, \cdots, C_s^s$, (where C_k^s is the number of combinations of s things taken k at a time.

[12] The exact values of Model III could be obtained by multiplying the corresponding items of the basic series by the exact values of the function $10^4 \exp \{-\frac{1}{2}(0.1t)^2\}/\sqrt{2\pi}$, for integral values of t. This function was the basis of obtaining the 4th differences of Model III. Approximate values of Model III were found by using a set of weights composed of 95 numbers corresponding to the values of the above function for integral values of t from -47 to $+47$, with the numbers less than 1 rounded off to the nearest tenth and numbers greater than 1 to whole units. The numbers of the basic series were written on a ribbon which we slid along the column of weights. Inasmuch as the basic series consisted of zeros and ones, all of the computations were plain additions. For Model IVa, a ribbon with holes in the place of unities was constructed.

times the 4th differences of the numbers of Model III from the 7th to the 97th.[13]

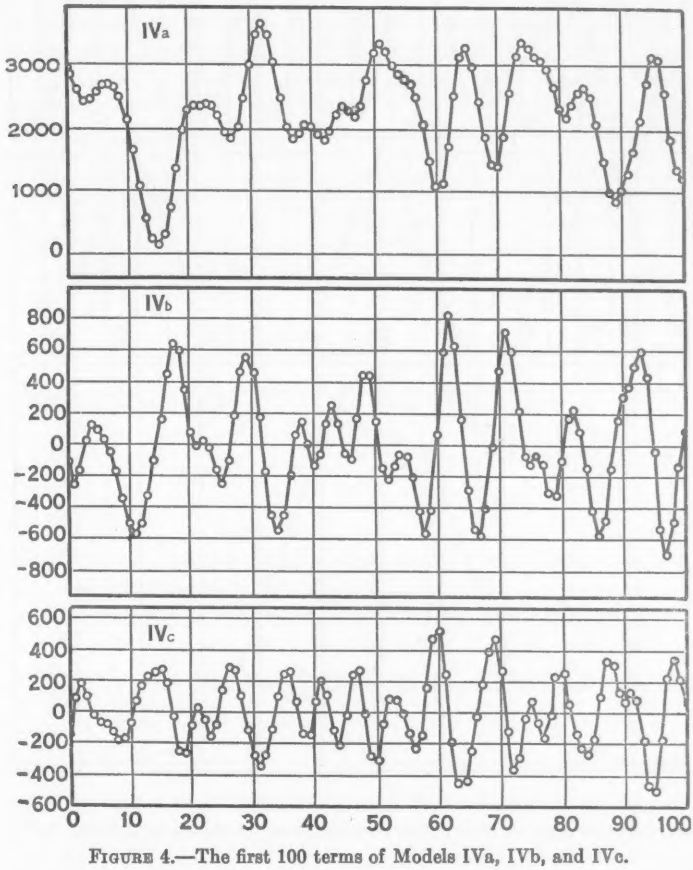


FIGURE 4.—The first 100 terms of Models IVa, IVb, and IVc.

We could not be satisfied by a smaller number of models because it was necessary to observe their various properties and to have illustra-

[13] For the calculation of these differences the accuracy with which we determined the items of Model III was not sufficient, so the following method was used: It is easy to see that the nth order differences of the items of the series obtained by scheme (1) are equivalent to those computed by the same scheme but applying weights equal to the differences of the original weights (keeping in mind that the series of original weights is extended at both ends with zeros).

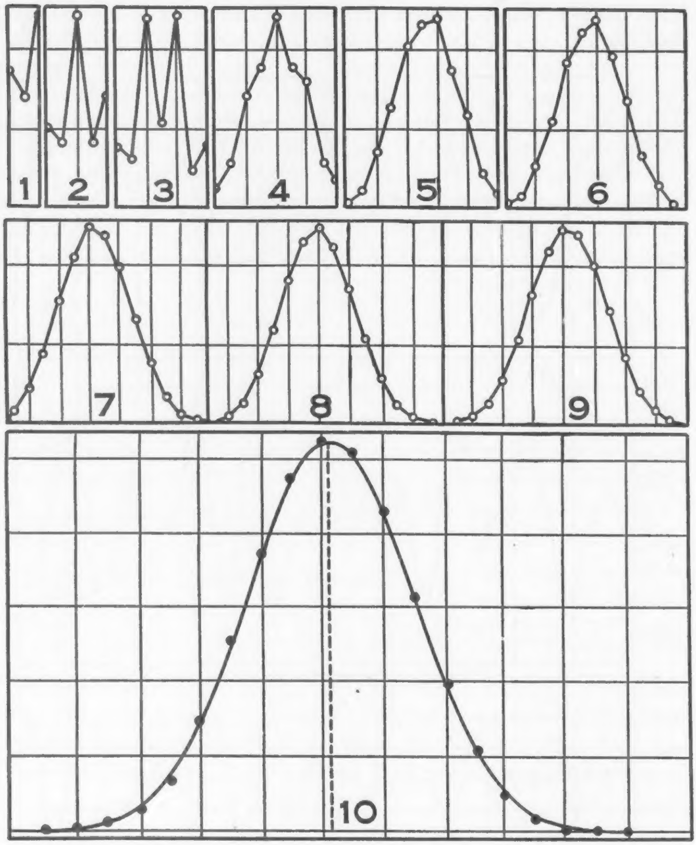


FIGURE 5.—An example of the crossing of random weights. The weights of the causes for each of 10 successive summations are shown, i.e., $A_k'^{(1)}$, $A_k'^{(2)}$, $\cdots$, $A_k'^{(10)}$. See Appendix, Section 1.

With the help of S. Pineto's *Tables de logarithmes vulgaires a dix decimales*, St. Petersburg, 1871, the values of the function

$$\exp\left\{-\tfrac{1}{2}(0.1t)^2\right\}/\sqrt{2\pi}$$

were obtained to ten decimal places for integral values of t from 0 to 44; this series was completed by using Sheppard's tables, and the differences of the entire series up to and including the 4th differences were taken. Multiplying the latter by 10^8 and expressing the result in integers, we obtained weights with the help of which—and by using scheme (1)—the values of $10^4\,\Delta^4 y_{III}$ were obtained from the second basic series.

tions for the elucidation of the different aspects of the problem. We could not aspire to imitate nature in forming a set of weights; still, in the course of the work, we have come across an exceptionally curious circumstance. First, each multifold simple summation of n items at a time gives a set of weights which approaches the Gaussian curve as a limit. In the Appendix, Section 1, there is given the instance of a tenfold summation of three items at a time with the weights chosen absolutely at random for each successive summation. The ten con-

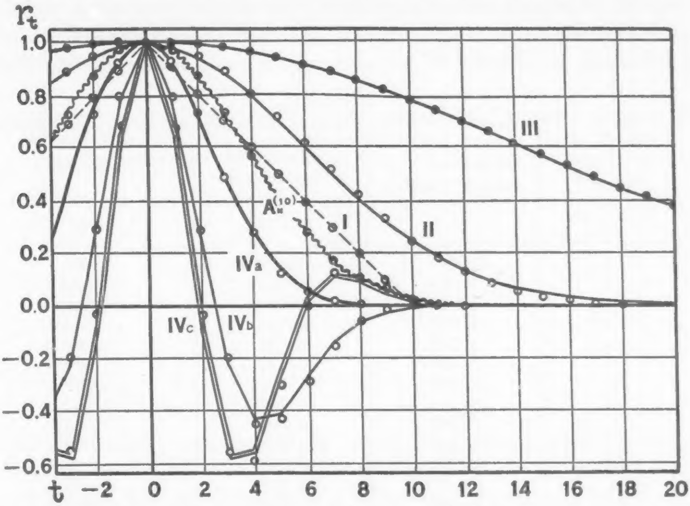


FIGURE 6.—oooo The correlational functions of Models I–IVc, and of the scheme of the crossing of the chance weights ($A_k'^{(10)}$).

Corresponding Gaussian curves and the reduced differences of the ordinates of the Gaussian curve.

secutive sets of weights are depicted in Figure 5. It is easily seen how they gradually become more and more like the Gaussian curve, and for the tenth summation the weights approach the Gaussian curve very closely.

This is far from being a chance result. From further considerations (Appendix, Section 1) we find that we have here actually encountered a law which, under certain conditions, must necessarily realize itself in the chaos of random entanglements and crossings of endless numbers of series of causes and consequences. The problem is specially important for the reason that the correlational function of a derived series is defined entirely by the respective weights-function. It is possi-

ble to prove (see Appendix, Sections 1, 2, and 3) that if the series of weights follows the Gaussian curve, the correlational function of the resulting consequence series is capable of being expressed by a similar curve with a greater or smaller degree of approximation. For the series of consequences proportionate to the increments of the cause—that is, the differences of order k of the series of causes—the correlational function can be represented by the series of the differences (of order $2k$) of the ordinates of the Gaussian curve. It could not be by chance that the correlational function of all of our models, with the exception of the most elementary one (Model I), belong to one of the two types mentioned (see Figure 6). No exception is found in the correlational

TABLE 1

Distance between terms	Correlation coefficients with random weights	Ordinates of Gaussian curve	Differences
t	r_t	R_t	$r_t - R_t$
0	1.000	1.000	0.000
1	0.965	0.965	0.000
2	0.868	0.866	+0.002
3	0.727	0.723	+0.004
4	0.567	0.562	+0.005
5	0.410	0.407	+0.003
6	0.275	0.274	+0.001
7	0.171	0.171	0.000
8	0.097	0.100	−0.003
9	0.051	0.054	−0.003
10	0.024	0.027	−0.003
11	0.011	0.013	−0.002
12	0.004	0.006	−0.002
13	0.001	0.002	−0.001
14	0.000	0.001	−0.001

function for the series of consequences of the 10th order obtained in the course of the crossing of the random weights in the example mentioned above. The values of these correlation coefficients (r_t), together with the ordinates of the corresponding Gaussian curve (R_t), are given in Table 1 (for the calculation see Appendix, Section 2).

III. THE UNDULATORY CHARACTER OF CHANCE SERIES; GRADUALITY AND FLUENCY AS TENDENCIES

Our models, representing several sets of experiments, give an inductive proof of our first thesis, namely, *that the summation of random causes may be the source of cyclic, or undulatory processes.*[14] It is, however

[14] The definition of the business cycle as being a process (not necessarily periodic) characterized by successive rises and falls, is given by *W. C. Mitchell* in Introduction to *W. L. Thorp, Business Annals*, New York, 1926, pp. 32–33.

not difficult to determine the reason why it must be so inevitably. We shall first observe a series of independent values of a random variable. If, for the sake of simplicity, we assume that the distribution of probabilities does not change, then, for the entire series, there will exist a certain horizontal level such that the probabilities of obtaining a value either above or below it would be equal. The probability that a value, which has just passed from the positive deviation region to the negative, will remain below at the subsequent trial is $\frac{1}{2}$; the probability that it will remain below two times in succession is $\frac{1}{4}$; three times $\frac{1}{8}$; and so on. Thus the probability that the values will remain for a long time above the level or below the level is quite negligible. It is, therefore, practically certain that, for a somewhat long series, the values will pass many times from the positive deviations to the negative and vice versa. Let us designate as a *half-wave* a portion of the series in which the deviation does not change sign. Thus, for 1000 numbers of the third basic series we find 540 half-waves (instead of the theoretically expected 500). Taking from this number the first 512 half-waves we find among them a number of half-waves of the length 1,2, etc. In Table 2 the actual (n_i') and theoretical[15] (n_i) frequencies for half-waves of various lengths are shown. That the observed series is consistent with the theoretical series can be found by the calculation of the χ^2 criterion of goodness of fit.[16]

TABLE 2

Length of half-wave	Actual frequency	Theoretical frequency
i	n'_i	n_i
1	261	256
2	137	128
3	65	64
4	29	32
5	14	16
6	4	8
7	1	4
8 and more	1	4
Total	512	512

If a variable can have more than two values and if, in a certain interval of a more or less considerable length, it happens to remain above

[15] L. von Bortkiewicz, *Die Iterationen*, 1917, Formel 75, p. 99.

[16] We find, indeed,

$$\chi^2 = \sum \frac{(n_i' - n_i)^2}{n_i} = 7.78,$$

the corresponding probability being $P = 0.35$; see *Tables for Statisticians and Biometricians*, ed. by K. Pearson, Part I, Table XII.

(or below) its general level, then in that interval it will have a temporary level about which it almost certainly will oscillate. Thus on the waves of one order there appear superimposed waves of another order.

The unconnected random waves are usually called irregular zigzags. A correlation between the items of a series deprives the waves of this characteristic and introduces into their rising and falling movements an element of *graduality*. In order to make the reasoning more concrete, let us consider a series obtained from an incoherent series by means of a ten-item moving summation. Our Model I will be used as the example. Any items of this model separated from each other by more than 9 intervals (as, for example, the values $y_0, y_{10}, y_{20}, \cdots$) are not correlated with each other and consequently form waves of the above considered type, i.e., irregular zigzags. But if we consider the entire series, we shall certainly find gradual transitions from the maximum point of a wave to its minimum and vice versa, since the correlation between neighboring items of the series makes small differences between them more probable than large ones. This we find to be true for all of our models.

We must distinguish between the *graduality* of the transitions and their *fluency*. We could speak about the absence of the latter property if a state of things existed where there would be an equal probability for either a rise or a fall after a rise as well as after a fall. If fluency were missing we should obtain waves covered by zigzags such as we find in Model I (see Figure 3).

For example, we have for Model I,

$$\begin{aligned} y_0 &= 5 + x_0 + x_1 + x_2 + \cdots + x_9, \\ y_1 &= 5 + x_1 + x_2 + \cdots + x_9 + x_{10}, \\ y_2 &= 5 + x_2 + \cdots + x_9 + x_{10} + x_{11}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

from which we obtain

$$\begin{aligned} \Delta y_0 &= y_1 - y_0 = x_{10} - x_0, \\ \Delta y_1 &= y_2 - y_1 = x_{11} - x_1, \\ &\cdots\cdots\cdots\cdots. \end{aligned}$$

Thus we see that the adjacent first differences do not have any causes in common, and hence are not correlated. The same applies to differences which are further apart, with the exception of such as $y_1 - y_0 = x_{10} - x_0$ and $y_{11} - y_{10} = x_{20} - x_{10}$. The series of differences is almost incoherent and hence the waves will be covered by chaotically irregular zigzags such as we find in Model I.

Let us assume further that adjacent differences are positively correlated. Then, in all probability, after a rise another rise will occur, after a fall a further fall; a steep rise will have the *tendency* to continue

with the same steepness, a moderate one with the same moderateness. So small sections of a wave will tend to be straight lines; and the greater the coefficient of correlation between adjacent differences the closer the sections approximate straight lines.[17]

Correlation between second differences plays an analogous role. The greater this correlation coefficient, the greater the tendency toward the preservation of the constancy of the second differences. Over more or less considerably long intervals a series with approximately constant second differences will tend to approximate a second-degree parabola as all "good" curves do. In Table 3 are given, for Models, I, II, and III, the values of the correlation coefficients between the adjacent items of the series (r_1), between the adjacent first differences ($r_1^{(1,1)}$), and between the adjacent second differences ($r_1^{(2,2)}$). The coefficients were calculated by the formulas of the Appendix, Section 1. As we go from the first basic series to Model I and then to Models II and III, we find progressive changes in their graphic appearance (see Figures, 3, 2, 8, and 11 respectively). These changes are produced at first by the introduction and then by the growth of graduality and of fluency in the movements of the respective chance waves. The growth of the degree of correlation between items (or between their differences) as we go from the first basic series to Model I, etc. (see Table 3) corresponds to the changes in the graphic appearance of our series.

TABLE 3

Model	Coefficient of correlation between:		
	Terms	First differences	Second differences
	r_1	$r_1^{(1,1)}$	$r_1^{(2,2)}$
I	0.9	0.0	−0.5
II	0.985	0.85	0.0
III	0.9975	0.9925	0.9876

IV. EMPIRICAL EVIDENCE OF THE APPROXIMATE REGULARITY OF CHANCE WAVES

Our first thesis, that is, the demonstration of the possibility of the appearance of undulatory processes of a more or less fluent character as the result of the summation of random causes, may be considered

[17] The term *tendency* is used here in a strict sense. To each equation of regression (giving the value of the conditional mathematical expectation of a variable as a function of some other variable) there corresponds an approximate equation between the variables themselves. The closer to unity the absolute value of the coefficient of correlation lies, the greater is the probability that this functional relationship will be maintained within the limits of the desired accuracy; i.e., the stronger will be the *tendency*.

as practically proved. However, our second thesis, that is, the demonstration of *the approximate regularity of the waves*, offers considerably greater difficulties. Again we shall begin with the inductive method.

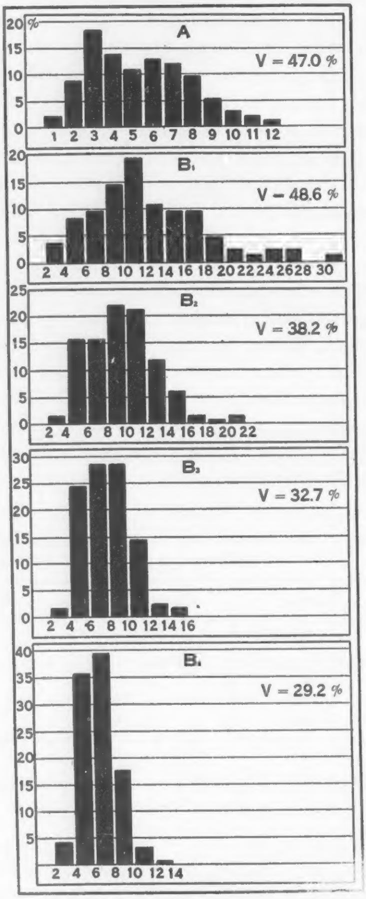


FIGURE 7.—The frequency distributions of the lengths of waves and half-waves: A, Business cycles of 12 countries, not including England (Mitchell); B_1 to B_4, Models II, IVa, IVb, and IVc respectively.

In Figure 2 are plotted 1000 points of Model II; a continuous line has been passed through them, which, because of the small scale, seems

to be a comparatively fluent curve. One can distinguish on the curve waves of different orders—even down to insignificant zigzags, of which a number are not apparent on the graph because of their minuteness. The maxima and minima having been listed, together with the length of their half-waves and amplitudes, we have found that, since an empirically descriptive point of view, in its very nature, permits only approximate solutions,[18] it was legitimate to draw a boundary between waves and *ripples:* maxima and minima with amplitudes of ten units or less being discarded as ripples. The remaining maxima and minima are indicated by arrows in Figure 2. The distribution of the lengths of the 83 half-waves for Model II is given graphically in Figure 7 (B_1). Figure 7 includes the distribution (A) of the lengths of 93 cycles of economic life for 12 countries outside of England, as given by Mitchell.[19] The coefficient of variation for the latter is 47.0%[20] as compared to 48.6% for Model II. Thus we find variation of the same degree in the two distributions. The distributions for Models IVa, IVb, and IVc are also shown in Figure 7. The average lengths of waves are 9.23, 7.36 and 6.15, while the coefficients of variation are 38.2%, 32.7% and 29.2% respectively. In general appearance these last three distributions are similar to the first two, although the last three have less variation, in spite of the fact that for Models IV, a, b, and c, the data are taken without discarding the ripples. Our models being based on some a priori schemes, it appears quite likely that some day it will be possible to calculate the mathematical expectation and variability of the distances between maxima and minima. In this respect, therefore, the chance waves in coherent series must be subject to some kind of regularity, the regularity of this type being observed even in the chaotic zigzags of purely random series.[21]

We are interested, however, in a different aspect of the problem. The attempt of Mitchell to deny the periodicity of business cycles is a result of his tendency to stick to a purely descriptive point of view. The means of description which he uses and which we tried to imitate for our models are far too crude. If we try to apply the same method to a sum of two or three sinusoids the result would be approximately the same. Those investigators of economic life are right who believe in their acumen and instinct and subscribe to at least an approximate correctness in the concept of the periodicity of business cycles. Let us

[18] Cf. E. Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, Halle a.d.S., 1922, § 74: *Deskriptive und exakte Wissenschaften*, p. 138–139.

[19] W. L. Thorp, *Business Annals*, Introduction by W. C. Mitchell, p. 58.

[20] *Ibid.*

[21] Cf. L. von. Bortkiewicz, *Die Iterationen*, 1917.

again examine Model II (Figure 2). In many places there are, apparently, large waves with massive outlines as well as smaller waves lying, as it were, over them; sometimes these are detached from them, sometimes they are almost completely merged into them. For example, at the beginning of Figure 2, three waves of nearly equal length are apparent, that is, from the first to the third minimum, from the third to the fifth, and from the fifth to the sixth. Upon these waves smaller

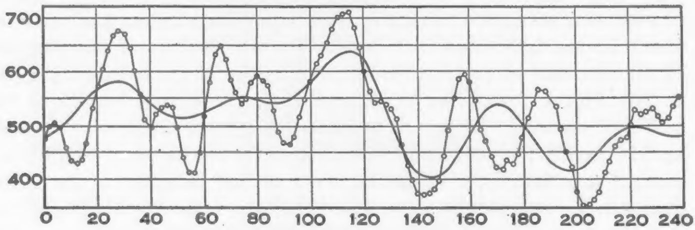


FIGURE 8.—o-o-o The first 120 even terms of Model II. ——— Sum of the first five harmonics of Fourier series: $y=518.14-20.98\cos(2\pi t/240)+50.02\sin(2\pi t/240)+17.30\cos(2\pi t/120)-3.16\sin(2\pi t/120)-10.93\cos(2\pi t/80)+35.66\sin(2\pi t/80)+17.18\cos(2\pi t/60)-21.92\sin(2\pi t/60)-38.53\cos(2\pi t/48)-3.65\sin(2\pi t/48)$.

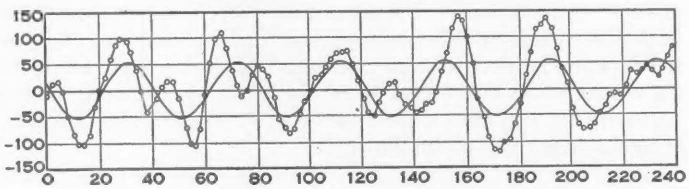


FIGURE 9.—o-o-o The deviations of Model II from the sum of the first five harmonics of Fourier series. ——— 6th sinusoid: $y=12.98\cos(2\pi t/40)-51.50\sin(2\pi t/40)$.

ones can be seen having also approximately equal dimensions. A careful examination of the graphs of our models will disclose to the reader a number of places where the approximate equality of the length of the waves is readily apparent. If we had a much shorter series, such as a series offered by the ordinary statistics of economic life with its small number of waves, we should be tempted to consider the sequence as strictly periodic, that is, as composed of a few regular harmonic fluctuations complicated by some insignificant casual fluctuations. For instance, let us consider two sections of Model II, lying one directly above the other in Figure 2, namely, the section from item 100 to

item 250 and the one from 600 to 750. The similarity between the waves in these sections is apparent.

The accuracy of the above deduction is limited by the imperfection of a visual impression. To eliminate this shortcoming, let us analyze one or two sections of our models harmonically by means of Fourier's analysis. This has been done for a section of 240 points of Model II and the 360 points of Model III. Because of the great fluency of these series it was sufficient to use only the even-numbered ordinates (i.e., 0, 2, $\cdots$ 238, and 0, 2, $\cdots$ 358, respectively), thus saving some computation. The results for the 120 points of Model II are shown in Figures 8, 9 and 10, those for the 180 points of Model III in Figures 11 and 12.

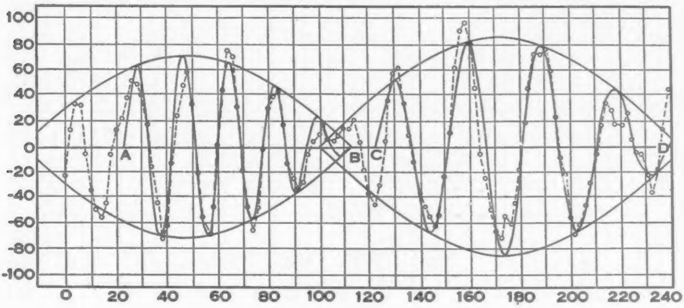


FIGURE 10.—o-o-o The deviations of Model II from the sum of six harmonics.

$$A\text{———}B\colon y_I = 71 \sin \frac{2\pi}{264}(t + 18) \sin \frac{2\pi}{18}(t - 24).$$

$$C\text{———}D\colon y_{II}\ 85 \sin \frac{2\pi}{288}(t - 100) \sin \frac{2\pi}{14\frac{2}{3}}(t - 122\tfrac{1}{3})$$

First let us consider Model II. In Figure 8 the sum of the first five sinusoids of the Fourier series are shown, while in Figure 9 the deviations from that sum are shown together with the sixth sinusoid. It is known, of course, that practically any given curve can be represented by a sum of a series of sinusoids provided a large enough number of terms is taken. It is not for every empirical series, however, that we can obtain such a significant correspondence and such a sharply expressed periodicity with a comparatively small number of harmonics. The approximately regular waves which were apparent even in the crude series are much more distinct now when they are isolated by deducting the sum of the first five harmonics. Of course, we cannot assert that the rest of Model II would follow the same periodicity, but, for our purposes, it is sufficient that successive waves should maintain an approxi-

mate equality of length for six periods. This hardly can be considered to be a chance occurrence; the explanation of such an effect must be found in the mechanism of the connection of the random values.

The deviations from the sum of six harmonics are plotted in Figure 10 together with the corresponding fluent curves. These curves are obtained as interference waves of two sinusoids with equal amplitudes and approximately equal periods. In other words, such a curve can be represented as the product of two sinusoids or as a sinusoid with an amplitude also changing along a sinusoid. These *bending sinusoids*

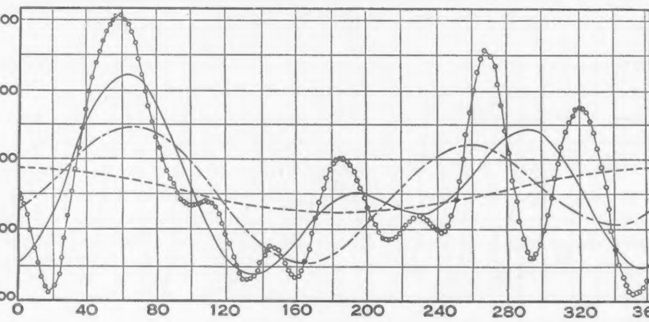


FIGURE 11.—o-o-o The first 180 even terms of Model III. —$y_I=554.8+31.79\cos(2\pi t/360)+3.40\sin(2\pi t/360)$. ----- $y_{II}=y_I-58.82\cos(2\pi t/180)+46.63\sin(2\pi t/180)$. ——— $y_{III}=y_{II}-75.36\cos(2\pi t/120)+0.61\sin(2\pi t/120)$.

separate on the graph the regions which place our empirical series in a definite *regime*.[22] Over a large part of the first region the regime is maintained for three or four periods with a correspondence that is much greater than could reasonably be expected between an analytical curve and a random series. At the beginning and end of a region the regime is broken. The point where a bending sinusoid cuts the axis of abscissas is the *critical point*. After this point a regime is replaced by another regime of the same type, but having different parameters. Throughout the greater part of the second region, as in the first, the regime is quite well sustained.[23]

[22] The term *regime* has been borrowed for the purposes of theoretical statistics from hydrography by N. S. Tchetverikov. See his work: "Relation of the Price of Wheat to the Size of the Crop," *The Problems of Economic Conditions*, Vol. 1, Issue 1, Moscow, 1925, p. 83.

[23] The parameters of a regime,

$$y = A\sin\left[(360°/L)(x-a)\right]\sin\left[(360°/l)(x-b)\right]$$

are easy to determine by means of graphical construction after a few trials. It is also possible to make corrections, using the method of least squares, but in our case we did not think it necessary.

If a result like the foregoing is not due to chance, a much better proof could be expected from an analysis of Model III for which the correlation between the elements is greater than for Model II. In Figure 11 the even-numbered points from 0 to 358 of Model III are plotted together with the first harmonic of the Fourier series, the sum of the first two, and the sum of the first three sinusoids. Instead of the six sinusoids needed for Model II, only three are here necessary for our purposes. The deviations from these are shown in Figure 12. Three regions are apparent with a change of regimes at the critical

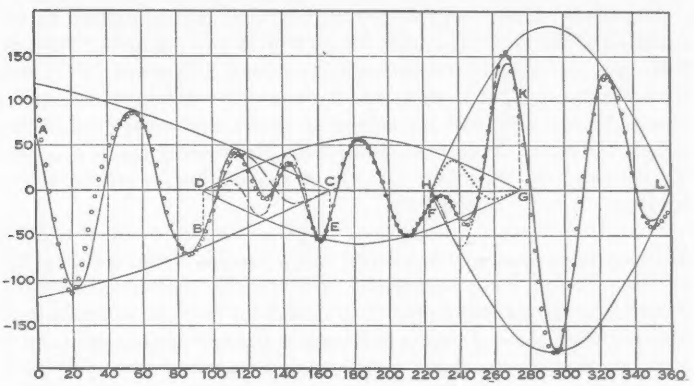


FIGURE 12.—oooo Deviations of Model III from the sum of the first three harmonics. A——B—-—C, Regime I: $y_I = 136 \sin [2\pi(t-167)/960] \sin [2\pi(t-39)/64]$. D—·—·—$E$, Regime II′: $y_{II}' = 58 \sin [2\pi(t-94)/360] \sin [2\pi(t-98)/36]$. E——F····G, Regime II″: $y_{II}'' = 58 \sin [2\pi(t-94)/360] \sin [2\pi(t-170)/54.4]$. H—··—··—K——L, Regime III: $y_{III} = 182 \sin [2\pi(t-222)/276] \sin [2\pi(t-250.6)/59.6]$. B——E; $y = y_I + y_{II}'$. F——K: $y = y_{II}'' + y_{III}$.

points. In addition we find one more regularity: to the overlapping parts of the said regions corresponds every time the partial superposition of the regimes, i.e., the algebraical addition of the respective curves.

Let us try now to summarize our observations in the following tentative and hypothetical manner:

The summation of random causes generates a cyclical series which tends to imitate for a number of cycles a harmonic series of a relatively small number of sine curves. After a more or less considerable number of periods every regime becomes disarranged, the transition to another regime occurring sometimes rather gradually, sometimes more or less abruptly, around certain critical points.

V. THE TENDENCY TO SINUSOIDAL FORM

In addition to the tendencies towards graduality and fluency (that is towards linear and parabolic forms for small sections) we find a third tendency, namely, the tendency toward a sinusoidal form.

Let $y_i, y_{i+1}, y_{i+2}, \cdots$ be the ordinates of a sinusoid. Then it is always true that

$$\Delta^2 y_i = -a y_{i+1}, \tag{2}$$

where $\Delta^2 y_i = (y_{i+2} - y_{i+1}) - (y_{i+1} - y_i)$, that is, the ith second difference of the series.

Conversely, it can easily be proved that the function defined by an equation of the form (2) in case $0 < a < 4$ must be a sinusoid.[24] Now, if there is a high correlation between the second differences ($\Delta^2 y_i$) and the ordinates (y_{i+1}) of a series, then equation (2) will be approximately true and there will exist a tendency toward a sinusoidal form in the series. The closer the correlation coefficient between $\Delta^2 y_i$ and y_{i+1}, denoted by us by $r_1^{(2,0)}$, is to -1, the more pronounced (or strong) is the tendency to a sinusoidal form.

A tendency toward either linear or parabolic forms cannot appear in a very large section of a coherent series because it would disrupt its cyclic character. The accumulation of deviations necessarily destroys every linear or parabolic regime even though the respective correlations are very high. After a regime is disrupted the new section will have a new, let us say a parabolic, regime (i.e., a regime of parabolas with different parameters). This process continues throughout the entire series, so that each coherent series of the type considered here is patched together out of a number of parabolas with variable parameters whose variations generally cannot be foreseen.

A sinusoidal regime is also bound to disrupt gradually, this being a property which distinguishes every tendency from an exact law. But under favorable conditions the sinusoidal tendency can be maintained over a number of waves without contradicting the basic property of a coherent series. In order to obtain a result of this kind it is necessary that the respective correlations be sufficiently high. But, as a matter of fact, $r_1^{(2,0)}$ for Model II is approximately the same as for Model I (-0.315 and -0.316), while Model III with its great smoothness has an $r_1^{(2,0)}$ less than that of Model IVa (-0.578 as compared to -0.599). It seems, however, to be very probable that this criterion is insufficient just because we have to deal here not with one sinusoid but with a whole series of sinusoids having different periods. Equation (2), of

[24] The condition $0 < a < 4$ is always satisfied in our case since $a = 2\ (1 - r_1)$ where r_1 is a correlation coefficient between the adjacent terms of the series (between y_i and y_{i+1}). See Appendix, Section 4.

course, is true only for a single sinusoid and cannot be applied to a sum of sinusoids.

To find an instance more apt to illustrate the tendency in question, let us consider the differences of various orders for Model III, the series best adapted for such purposes. If a curve is represented by a sum of sinusoids, then the differences of all orders are sums of sinusoids having waves of the same periods as the curve. The higher the order of the difference, the more pronounced are the shorter periods, since the differencing process weights the shorter periods as against the longer ones. Thus, by applying the formulas of the Appendix, Section 1, we find

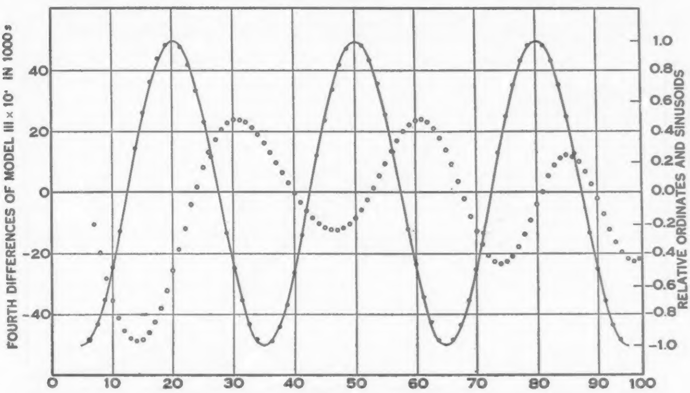


FIGURE 13.—oooo The fourth differences of Model III$\times 10^4$. —·—·— Their relative ordinates, together with the sinusoid.

that the coefficients of correlation, $r_1^{(2,0)}$, for Model III itself, and for its first, second, third, and fourth differences, are -0.5781, -0.7756, -0.8462, -0.8830, and -0.9057, respectively. The following considerations will show us to what extent the simple sinusoidal regime is maintained at least over small portions of the last series.

Let us determine the highest (or lowest) point of a typical wave (A or B, respectively, in Figure 14) as the apex of a second-degree parabola which passes through the three highest (or the three lowest) points of the wave. Then, let us draw a horizontal line bisecting the distance between the highest and lowest points of the wave. Further, let us denote the point where this horizontal line crosses the straight line joining the two points between which the horizontal line passes as C. This point divides AB into two quarter-waves, AC and BC. For each of these, let us make the following construction: Dividing the base line DC (D having the same abscissa as A) into six equal parts, we obtain seven

points corresponding to 0°, 15°, 30°, 45°, 60°, 75°, and 90°. At the five central points construct perpendiculars and extend them to the parabola fitted to the three empirical points (interpolated according to Newton's formula). These perpendiculars are the ordinates of an empirical half-wave and, if we divide through by the maximum ordinate AD, we obtain the relative ordinates y_{15}, y_{30}, y_{45}, y_{60}, and y_{75}. If our wave is a sinusoid, these relative ordinates will equal the sines of 15°, 30°, 45°, 60°, and 75°, respectively. The empirical relative ordinates for the 12 quarter-waves of $\Delta^4 y_{III}$ are shown by black dots around the regular sinusoid of Figure 13, while the relative ordinates of the first, second, etc., quarters of every empirical wave are shown on the first,

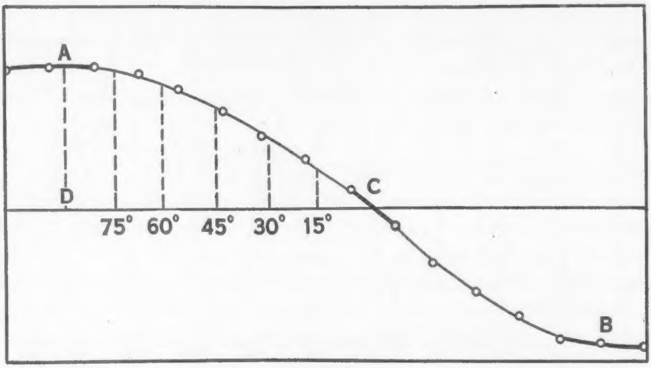


FIGURE 14.—A Scheme for Calculation of the Relative Ordinates.

second, etc., quarters of the sinusoid. The points can hardly be distinguished from the curve. Thus the tendency to a sinusoidal form is shown rather distinctly. If we compute the arithmetic averages of the relative ordinates having the same abscissa (e.g., $\bar{y}_{15} = 1/12(y_{15}^{(1)} + y_{15}^{(2)} + \cdots + y_{15}^{(12)})$ and compare them with the corresponding sines (e.g., sin 15°), we shall see that the deviations are less than $\frac{1}{2}$ in the second decimal place (see Table 4). The agreement is, therefore, close enough

TABLE 4

Phase-angle (α)	15°	30°	45°	60°	75°
$\bar{y}_\alpha$	0.258	0.496	0.703	0.863	0.964
Sin α	0.259	0.500	0.707	0.866	0.966
Deviations	−0.001	−0.004	−0.004	−0.003	−0.002

to be considered as the clear manifestation of the tendency toward a sinusoidal form, and thus displays once more the ability of chance waves to simulate regular harmonic oscillations.

VI. ON THE PSEUDO-PERIODIC CHARACTER OF THE EMPIRICAL CORRELATIONAL FUNCTION[25]

As a further illustration of the sinusoidal tendency, I shall consider here a chance series satisfying, to a rather high degree of approximation, the equation

$$\Delta^4 z_i - p\Delta^2 z_{i+1} - q z_{i+2} = 0, \tag{3}$$

corresponding, if treated as a precise one, to the sum of two sinusoids.

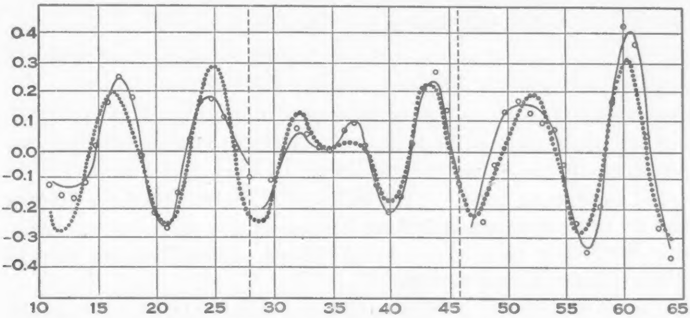


FIGURE 15.— o o o o The reduced empirical correlation function of Model IVc. ——— Sum of two sinusoids, separately for each of three intervals. ..... Sum of three sinusoids.

Let us denote the series in question by the symbols

$$\rho'_{11}, \rho'_{12}, \cdots, \rho'_{64},$$

the values of ρ'_t (see Figure 15) being given by the equation

$$\rho'_t = \sqrt{\frac{128 - t}{128}}\, \rho_t, \tag{4}$$

where ρ_t is the empirical correlation coefficient between the terms (y_i, y_{i+t}) of the series made up by 128 items of Model IVc. As the values of ρ_t have been calculated from very different numbers of items varying from 117 (that is, $128-11$) to 64 (that is, $128-64$), the reduction

[25] Eugen Slutzky, "On the Standard Error of the Correlation Coefficient in the Case of Homogenous Coherent Chance Series" (in Russian, with English summary). Transactions of the Conjuncture Institute, Vol. 2, 1929, pp. 94–98, 154.

by (4) has been thought useful in order to bring the respective standard deviations to approximate equality.

Before going further, the following remarks will be made. Let y_0, $y_1, \cdots$ be a stationary chance series. This implies that the mathematical expectation, $E(y_i)$, is a constant, that the standard deviation, σ, is a constant, and that the correlation coefficient, r_t, between y_i and y_{i+t}, is a function of t only. Then, putting, without loss of generality, $E(y_i)=0$, we shall have $\sigma^2=E(y_i^2)$ and $r_t=E(y_iy_{i+t})/\sigma^2$. This being the *theoretical* correlation coefficient, let us suppose that $r_t=0$ if $t>\omega$. Then the correlation coefficient, r'_u, between the empirical correlation coefficients, ρ_t and ρ_{t+u}, will be given by the equation

$$r'_u = r(\rho_t, \rho_{t+u}) = \frac{\sum_{-\omega}^{\omega} r_t r_{t+u}}{\sum_{-\omega}^{\omega} r_t^2}, \tag{5}$$

this formula being approximately correct if it be supposed (1) that ρ_t and ρ_{t+u} are calculated from the same number of values, n; (2) that n is sufficiently large; and (3) that $t>2\omega$, $n-t>\omega$, and $u>0$.[26]

Let us suppose now that the values of r'_u calculated from (5) may be held to be approximately true for the series of the reduced correlation coefficients $(\rho'_{11}, \rho'_{12}, \cdots \rho'_{64})$ defined above. Then we have to consider the following problem:

The series given (ρ'_t) being a chance series, there can exist no periodicity in the strict sense of the word. Its cyclical character being, however, obvious (see Figure 15), it may be asked whether the law of its composition from simple harmonics cannot be detected when its correlational function is known.

Let us try to solve this problem to the first approximation by supposing that our series can be duly approximated by the sum of two sinusoids with constant periods and varying amplitudes and phases. To this end, let us find the parameters of the regression equation which can be written in the form

$$\Delta^4\rho'_u = p\Delta^2\rho'_{u+1} + q\rho'_{u+2} + \epsilon, \tag{6}$$

ϵ being the "error," and p and q being determined by the method of least squares. If we denote the correlation coefficients between the pairs of values

$$(\Delta^4\rho'_u, \Delta^2\rho'_{u+1}), (\Delta^4\rho'_u, \rho'_{u+2}), (\Delta^2\rho'_{u+1}, \rho'_{u+2}),$$

by r_{12}, r_{13}, r_{23} respectively, then, using the formulae (43)–(44) (Appendix, Section 1), we obtain

[26] Cf. Slutzky, *loc. cit.* in note 25, pp. 91–94.

$$(7) \qquad r_{12} = \frac{\Delta^6 r'_{-3}}{\sqrt{\Delta^8 r'_{-4} \Delta^4 r'_{-2}}}\,; \quad r_{13} = \frac{\Delta^4 r'_{-2}}{\sqrt{\Delta^8 r'_{-4}}}\,; \quad r_{23} = \frac{\Delta^2 r'_{-1}}{\sqrt{\Delta^4 r'_{-2}}},$$

where r'_{-u} $(=r'_u)$ is the correlation coefficient defined by the equation (5). These values are

$$r'_{-4} = r'_4 = -\ 0.761{,}874{,}30,$$
$$r'_{-3} = r'_3 = -\ 0.618{,}465{,}96,$$
$$r'_{-2} = r'_2 = -\ 0.013{,}793{,}10,$$
$$r'_{-1} = r'_1 = -\ 0.689{,}655{,}17,$$

whence, using (7),

$$r_{12} = 0.945{,}847, \quad r_{13} = 0.760{,}844, \quad r_{23} = -\ 0.919{,}999.$$

Then, by the well-known formula of linear regression, we obtain

$$p = -\ 1.419{,}386, \quad q = -\ 0.425{,}828,$$

the multiple correlation coefficient, between $\Delta^4\rho_u$ on the one hand, and $\Delta^2\rho_{u+1}$ and ρ_{u+2} on the other, being $r_{1.23}=0.986$. The correlation is thus very high and so it is quite reasonable to omit ϵ in (6) and to treat the resulting approximate equation according to the rules of the calculus of finite differences. We find thus that the solution of this equation is the sum of two sinusoids with the periods

$$L_1 = 9.40, \quad L_2 = 6.04.$$

Let us, then, divide our series $(\rho'_{11}, \rho'_{12}, \cdots \rho'_{64})$ into three parts, of 18 items each, and let us find, for each part separately, two sinusoids with the periods $L_1=9, L_2=6$, these being the whole numbers nearest to the theoretical values just obtained. We find the results given in Table 5.

TABLE 5

	Part I		Part II		Part III	
	$L_1=9$	$L_2=6$	$L_1=9$	$L_2=6$	$L_1=9$	$L_2=6$
Amplitude	0.19790	0.07820	0.12614	0.12527	0.29123	0.13340
Phase	231°37′	52°39′	218°2′	295°2′	270°52′	2°33′

A glance at Fig. 15 shows that the theoretical curves fit the empirical points very satisfactorily and it seems fairly certain that, if we had included one or two sinusoids more, we could have obtained a quite satisfactory fit, even if treating our empirical series as a whole. This can be proved by the fact that the sum of three sinusoids

$$\begin{aligned} z = 0.1893 \sin\ &[(360°/8.80)t - 47°3'] \\ &+ 0.1000 \sin\ [(360°/7.14)t + 168°24'] \\ &+ 0.0794 \sin\ [(360°/5.87)t - 76°3'], \end{aligned}$$

though found by rather a rough graphical estimate, fits our empirical curve in a fairly satisfactory manner.[27]

VII. THE LAW OF THE SINUSOIDAL LIMIT

Many tendencies dealt with rather empirically in the preceding discussion will be more clearly understood, and their significance more fully appreciated, if we take into consideration the following propositions.[28]

THEOREM A: (*The Law of the Sinusoidal Limit*)

Let $y_1, y_2, \cdots$ *be a chance series fulfilling the conditions,*

$$E(y_i) = 0, \quad E(y_i^2) = \sigma^2 = f(n),$$

$$\frac{E(y_i y_{i+t})}{E(y_i^2)} = r_t = \phi(t, n),$$

where n is a parameter specifying the series as a whole, and $f(n)$ and $\phi(t, n)$ are independent of i. If, furthermore, the correlation coefficient, r_1, between y_i and y_{i+1}, satisfies the condition

$$|r_1| \leqq c < 1, \qquad (n \to \infty),$$

and the correlation coefficient between $\Delta^2 y_i$ and y_{i+1}, that is, ρ_1, is such that

$$\lim \rho_1 = -1, \qquad (n \to \infty),$$

then (1) ϵ and η being taken arbitrarily small and s arbitrarily large, there will exist a number, n_0, such that for every $n > n_0$, the probability, that the absolute values of the deviations of $y_i, y_{i+1}, \cdots y_{i+s}$ from a certain sinusoid will not exceed $\epsilon\sigma$, will be $>1-\eta$; (2) the period of this sinusoid will be determined by the equation

$$\cos(2\pi/L) = r_1;$$

(3) the number of the periods in the interval $(i, i+s)$ will be arbitrarily large provided s and n be taken large enough.

This proposition (for its proof see Appendix, Section 4) would be of no interest could we not give at least a single instance of a chance series satisfying the conditions of Theorem A. This is done by

[27] The above illustration seems to throw some light on the difficulties connected with the idea of a correlation periodogram. Cf. Dinsmore Alter, "A Group or Correlation Periodogram," etc., *Monthly Weather Review*, Vol. 55, No. 6, June, 1927, pp. 263–266; Sir Gilbert Walker, "On Periodicity in Series of Related Terms," *Proc. Royal Soc.*, Ser. A, Vol. 131, No. A 818, 1931, pp. 518–532.

[28] E. Slutsky, "Sur un théorème limite relatif aux séries des quantités éventuelles," *Comptes Rendus*, Paris, t. 185, seance du 4 Juli, 1927, p. 169.

THEOREM B: Let $x_1, x_2, \cdots$ be a random series fulfilling the conditions

$$E(x_i) = 0,\ E(x_i^2) = \sigma_x^2 = \text{const.},\ E(x_i x_j) = 0, \qquad (i \neq j).$$

Now, if we put

$$x_i^{(1)} = x_i + x_{i-1},\ x_i^{(2)} = x_i^{(1)} + x_{i-1}^{(1)},\ \cdots,$$

$$x_i^{(n)} = x_i^{(n-1)} + x_{i-1}^{(n-1)},$$

and

$$y_i = \Delta^m x_i^{(n)},$$

then the series $y_1, y_2, \cdots$ will tend to obey the law of the sinusoidal limit, provided m and n be increasing indefinitely and m/n = constant (for the proof see Appendix, Section 4).

Both propositions can be generalized to the case of a chance series practically coinciding, not with one sinusoid, but with the sum of a certain number of sinusoids.[29] In every case, however, the practical coincidence (and it is a very essential character of the series under consideration) does not extend itself to the series as a whole, the respective sinusoids of closest fit being different for different partial series. This is plainly evident for the chance series of Theorem *B*, for, *s* being arbitrarily large and *n* and *m* being sufficiently large, the values y_i and y_{i+t} will be wholly independent of each other as soon as $t > m + n + 1$, whence it follows that the phases and the amplitudes of the sinusoids practically coincident with the partial series, $y_i, y_{i+1}, \cdots y_{i+s}$, and $y_{i+t+s}, y_{i+t+s+1}, \cdots y_{i+t+2s}$, respectively, will also be independent of each other provided $t > m+n+1$.

The following considerations will show us the same problem from a somewhat different standpoint. Let us suppose a certain mechanism is being subjected to damped vibrations of a periodic character and to casual disturbances accumulating energy just sufficient to counterbalance the damping.[30] Then the movement of the system could be regarded as consisting of the two parts: of the vibrations determined by the initial conditions at some given moment, and of the vibrations generated by the disturbances that have occurred since. As soon as the first part has been nearly extinguished by the damping process after due lapse of time, the actual vibrations will be reduced practically to the second part, that is, to the accumulated consequences of the chance

[29] V. Romanovsky, "Sur la loi sinusoidale limite," *Rend. d. Circ. mat. di Palermo*, Vol. 56, Fasc. 1, 1932, pp. 82–111; V. Romanovsky, "Sur une generalisation de la loi sinusoidal limite," *Rend d. Circ. mat. di Palermo*, Vol. 57, Fasc. 1, pp. 130–136; cf. Sir Gilbert Walker, *op. cit.*, in note 27.

[30] Cf. G. Udney Yule, "On a Method of Investigating Periodicities in Disturbed Series with Special Reference to Wolfer's Sunspot Numbers," *Phil. Trans. Roy. Soc. of London*, Ser. A. Vol. 226, 1927, pp. 267–298.

causes. The latter, after a due time, being again reduced to a value not different practically from zero, the vibrations will consist of the disturbances accumulated during the second interval of time and so on. It is evident that the vibrations ultimately will have the character of a chance function, the described process being a particular instance of the summation of random causes. Should the disturbances be small enough, there would exist an arbitrarily large, but finite, number, L_0, such that the resulting process would be practically coincident, in every interval of the length, $L \leq L_0$, with a certain periodic (or nearly periodic) function, obeying thus the law of the sinusoidal limit.

Analogous considerations may be applied to the motion of planetary, or star systems, the innumerable cosmic influences being considered as casual disturbances. The paths of the planets, if regarded during billions of years, should be considered, therefore, as chance functions, but if we do not wish to go beyond thousands of years their approximate representation must be taken as not casual.

The chance functions of the type just considered appearing on the one end of the scale, and the random functions on the other, there evidently must exist all possible intermediate gradations between these extremes. The ability of the coherent chance series to simulate the periodic, or the nearly periodic, functions, seems thus to be definitely demonstrated.

It remains for us to try to clear up theoretically the remarkable property of some specimens of chance series, which do not belong to the extreme classes of their type, of being approximately representable by a small number of sinusoids, over a shorter or longer interval.

It is well known that every empirical series consisting of a finite number of terms ($N=2n$ or $2n+1$) can be represented precisely by a finite Fourier series, that is by the sum of a finite number (n) of sinusoids. Further, it is plainly evident, the series under consideration being chance series and the coefficients of the Fourier expansion,

$$y_t = A_0 + \sum_1^n A_k \cos (2\pi kt/N) + \sum_1^n B_k \sin (2\pi kt/N),$$

that is, the values $A_0, A_1, \cdots A_n$, and $B_0, B_1, \cdots B_n$, being linear functions of $y_1, y_2, \cdots y_n$, that the variables A_k and B_k will also be chance variables. Their mathematical expectations, standard deviations, and the correlation coefficients between them can be easily obtained.[31] Denoting by R_k^2 the intensity of the kth harmonic, that is,

[31] E. Slutsky, "Alcuni applicazioni di coefficienti di Fourier al analizo di sequenze eventuali coherenti stazionarii," *Giorn. d. Istituto Italiano degli Attuari*, Vol. 5, No. 4, 1934; see also E. Slutsky, "Sur l'extension de la theorie de periodogrammes aux suites des quantités dependentes." *Compte Rendus*, t. 189, seance du 4 novembre, 1929, p. 722.

the square of its amplitude, we shall have

$$R_k^2 = A_k^2 + B_k^2,$$

and

$$(8) \quad \begin{aligned} E(R_k^2) &= (4\sigma_y^2/N)\left[1 + 2\sum_1^{N-1} r_t \cos(2\pi kt/N)\right] \\ &\quad - (8\sigma_y^2/N^2)\sum_1^{N-1} tr_t \cos(2\pi kt/N), \end{aligned}$$

whence, for the case of a random series, we obtain at once the formula of Schuster,

$$(9) \qquad E(R_k^2) = 4\sigma_y^2/N,$$

the probability distribution being the same in both cases,

$$(10) \qquad P(R_k^2 > Z^2) = \exp\left[-Z^2/E(R_k^2)\right].$$

Let us suppose the m intensities happening to have the largest values in some given case are those with the indices: $\alpha, \beta, \cdots \mu$ and let

$$\tfrac{1}{2}(R_\alpha^2 + R_\beta^2 + \cdots + R_\mu^2) = ps^2,$$

s^2 being the square of the empirical standard deviation and p the coefficient measuring the degree of approximation reached in the given case by means of m harmonics. By taking account of (8), (9), and (10), we see at once that, in the case of a random series, the indices $\alpha, \beta, \cdots \mu$ are able to assume any values with equal probability but that in the case of a coherent series those having the largest values of $E(R_k^2)$ will be the most probable. As half of the sum of the intensities is equal to the square of the empirical standard deviation (Parceval's theorem), it is but natural that the coherent chance series, in many cases at least, may be represented—the degree of approximation being the same—by a smaller number of harmonics than the random series.

It can be proved further (under suppositions of a not very restrictive character) that the correlation coefficients between the intensities belonging to the same interval, as well as between those belonging to the adjacent intervals, are quantities of the order $1/N^2$, and that the standard deviation of the intensity, σ_{R^2}, tends to be equal to its probable value, $E(R_k^2)$. Whence it is evident that the indices of the harmonics which happen to be the most suited for the representation of the series in a certain interval must also be practically independent of the indices of the "best" harmonics in adjacent intervals, the length of these intervals being sufficiently large. The larger the probable value of the intensity the larger also must be the extent of its casual variation. These are properties quite consistent with a considerable degree of regularity—as well as with the abrupt changes of the "regimes" de-

termined by studying the empirical series dealt with in the foregoing pages.

Appendix

MATHEMATICAL NOTES OF THE THEORY OF RANDOM WAVES

1. Let $x_0, x_1, \cdots x_i, \cdots$ be a random series, that is, a series of chance values independent of each other. Let this be our basic series and let it be considered as a model of incoherent series of random causes. Denoting by the symbol E the mathematical expectation, let us suppose that

$$E(x_i) = 0, \; E(x_i^2) = \sigma_x^2 = \text{const.}; \; E(x_i x_j) = 0, \qquad (i \neq j). \tag{1}$$

From the basic incoherent series of causes let us construct a coherent series of "consequences," $\cdots y_{i-2}, y_{i-1}, y_i, \cdots$, by the scheme

$$y_i = \sum_{k=0}^{n-1} A_k \, x_{i-k}, \tag{2}$$

where the quantities A_k are constants.[32] Then, by using (1) and (2), it can easily be shown that

$$E(y_i) = 0, \tag{3}$$

$$E(y_i^2) = \sigma_y^2 = \sigma_x^2 \sum_{k=0}^{n-1} A_k^2, \tag{4}$$

$$E(y_i y_{i+t}) = \sigma_x^2 \sum_{k=0}^{(n-1)-t} A_k A_{k+t}. \tag{5}$$

Since equations (4) and (5) do not depend on i, the coefficient of correlation between y_i and y_{i-t}, that is, r_t, is also independent of i, and we have

$$r_t = \frac{\sum_{k=0}^{(n-1)-t} A_k A_{k+t}}{\sum_{k=0}^{n-1} A_k^2}, \tag{6}$$

from which it immediately follows that

$$r_0 = 1; \; r_t = r_{-t}; \; r_t = 0, \; (t \geqq n). \tag{7}$$

The process of moving summation can be repeated. As before, let us take $\cdots x_{i-2}, x_{i-1}, x_i \cdots$ as the basic series underlying the conditions

[32] Cf. Prof. Birger Meidell's valuable investigation of the analogous cumulative processes in his paper, "Über periodische und angenäherte Beharrungszustände," *Skandinavisk Aktuarietidskrift*, 1926, p. 172.

(1). Then on performing an s-fold moving summation we obtain the following successive consequence series:

$$x_i^{(1)} = \sum_{k=0}^{n-1} \alpha_k^{(1)} x_{i-k}; \quad x_i^{(2)} = \sum_{k=0}^{n-1} \alpha_k^{(2)} x_{i-k}^{(1)}; \cdots \tag{8}$$
$$x_i^{(s)} = \sum_{k=0}^{n-1} \alpha_k^{(s)} x_{i-k}^{(s-1)}.$$

After the s-fold summation we have an expression of type (2) with $y_i = x_i^{(s)}$. Hence, if we take $n=2$ and $\alpha_k^{(j)}=1$, it can easily be shown that

$$y_i = x_i^{(s)} = \sum_{k=0}^{s} C_k^s x_{i-k}. \tag{9}$$

If we put

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}t^2\right), \tag{10}$$

we can, by the use of well-known transformations, obtain the approximate expression (which we write for an *even* s)

$$y_i = D \sum_{k=0}^{r} x_{i-k} \phi\left(\frac{k - s/2}{\sqrt{s/4}}\right), \tag{11}$$

D being a coefficient, the value of which need not concern us here.

It is very remarkable that a similar result will always be obtained for s sufficiently great, whatever be the weights used, provided it is supposed (1) that the weights are not negative, (2) that they remain constant at every given stage of the process, and (3) that the summation does not tend to degenerate into a mere repetition of the same values, which would be the case should all $\alpha_k^{(s)}$ but one tend to approach 0; (the sum of the weights is supposed, without loss of generality, to be constant).

To prove this, let us remark first that the result of the s-fold summation given by (8) can evidently be obtained by a similar s-fold summation with the weights

$$p_0^{(j)}, p_1^{(j)}, \cdots, p_{n-1}^{(j)}, \qquad (j = 1, 2, \cdots, s),$$

where

$$p_k^{(j)} = \frac{\alpha_k^{(j)}}{m_j}, \quad m_j = \sum_{k=0}^{n-1} \alpha_k^{(j)},$$

if we multiply the resulting weights by the proportionality factor $m_1 \cdot m_2 \cdot \cdots \cdot m_s$.

Now to prove our proposition we shall use the following analogy (kindly suggested to the author by Prof. A. Khinchin).

Let $z_1, z_2, \cdots z_s$ be a set of random variables whose possible values are 0, 1, 2, $\cdots$, $n-1$, the respective probabilities being $p_0^{(j)}$, $p_1^{(j)}$, $\cdots p_{n-1}^{(j)}$, $(j=1, 2, \cdots s)$. Then it is easy to see that the probability of the equation

$$k = z_1 + z_2 + \cdots + z_s \tag{12}$$

must be equal to the coefficient of x^k in the expansion of

$$\prod_{j=1}^{s} (p_0^{(j)} + p_1^{(j)}x + p_2^{(j)}x^2 + \cdots + p_{n-1}^{(j)}x^{n-1}). \tag{13}$$

On the other hand, it can be proved that the same coefficient, multiplied by $m_1 \cdot m_2 \cdot \cdots \cdot m_s$, will be equal to the coefficient A_k in the equation (2) obtained by an s-fold summation according to the scheme (8) with the weights

$$m_jp_0^{(j)},\ m_jp_1^{(j)},\ \cdots,\ m_jp_{n-1}^{(j)}, \qquad (j = 1, 2, \cdots, s).$$

This is easily seen for $s=2$ and the result can be generalized by mathematical induction from s to $s+1$.

This analogy leads us to the following considerations. Let us put

$$\begin{cases} a_j = E(z_j) = \sum_{k=0}^{n-1} kp_k^{(j)}, \\ b_j = E[(z_j - a_j)^2] = \sum_{k=0}^{n-1} (k - a_j)^2 p_k^{(j)}, \\ c_j = E[|z_j - a_j|^g] = \sum_{k=0}^{n-1} |k - a_j|^g p_k^{(j)}, \qquad (g > 2). \end{cases} \tag{14}$$

It is evident that $b_j=0$ only if every $(k-a_j)^2p_k^{(j)}=0$ for $k=0, 1, 2, \cdots (n-1)$, and that this is possible only when every $p_k^{(j)}$ but one is equal to zero, the exceptional p being 1, in which case the values $x_k^{(j)}$ are merely repetitions of the values $x_k^{(j-1)}$.

This case being excluded, we shall have, on the average at least,

$$(1/s)\sum_1^s b_i > \epsilon > 0; \tag{15}$$

whence

$$\frac{\left[\sum_1^s c_i\right]^2}{\left[\sum_1^s b_i\right]^g} < \frac{\left[(1/s)\ \sum_1^s c_i\right]^2}{s^{g-2}\epsilon^g} \to 0. \tag{16}$$

But this is the well known Liapounoff's condition, under which the

probability distribution of the sum $z_1+z_2+\cdots+z_s$, that is, the distribution of the coefficients A_k, tends to the normal law.[33]

Let us put, for instance,

$$\alpha_0^{(j)} = \alpha_1^{(j)} = \cdots = \alpha_{n-1}^{(j)} = 1, \qquad (j = 1, 2, \cdots, s).$$

Then we obtain

$$\begin{aligned} m_j &= \sum_{k=0}^{n-1} \alpha_k^{(j)} = n, \\ m_1 m_2 \cdots m_s &= n^s, \\ p_k^{(j)} &= 1/n, \end{aligned} \tag{17}$$

for

$$j = 1, 2, \cdots, s, \quad \text{and} \quad k = 0, 1, \cdots, n-1;$$

and

$$\begin{aligned} a_j &= E(z_j) = \sum_{k=0}^{n-1} k p_k^{(j)} = (n-1)/2, \\ b_j &= E[(z_j - a_j)^2] = (1/n)\left\{\sum_{k=0}^{n-1} k^2 - n a_j^2\right\} = (n^2-1)/12. \end{aligned} \tag{18}$$

Whence

$$k_0 = E(k) = E\left[\sum_1^s z_j\right] = s(n-1)/2, \tag{19}$$

and

$$\sigma_k = \sqrt{s b_j} = \sqrt{s(n^2-1)/12}. \tag{20}$$

As s tends towards ∞, the value of A_k will thus approach a limit, which enables us to write, for s large but finite, the following approximate equations:[34]

[33] It is evident that, since Liapounoff's theorem is a proposition about the limit properties of certain integrals and not of the individual ordinates, the above demonstration must be interpreted also in the same sense. For many cases, however, for example, in the case of the illustration below, the additional conditions are satisfied under which the values of the variables A_k themselves are tending toward the ordinates of the Gaussian curve.

Cf. Liapounoff, "Nouvelle forme du théorème sur la limite de probabilité," *Memoires de l'Academie de science de St.-Petersbourg*, serie 8, Vol. 12, No. 5.

R. von Mises, *Vorlesungen aus dem Gebiete der Angewandten Mathematik*, Bd. I—*Wahrscheinlichkeitsrechnung und ihre Anwendungen*, 1931, p. 200–212.

R. von Mises, "Generalizzazione di un teorema sulla probabilità della somma di un numero illimitato di variabili casuali," *Giornale dell'Istituto Italiano degli Attuari*, Anno 5, N4, p. 483–495.

[34] This result coincides with that given in the first edition of this memoir in 1927; it was supplied to the author by the courtesy of Prof. A. Khinchin who derived it by the application of the well-known Cauchy theorem to the evaluation of the coefficient of x^k in the expansion of

$$(1 + x + x^2 + \cdots + x^{n-1})^s.$$

I am sorry that the calculations are too long to be reproduced here.

$$(21) \qquad A_k \overset{a}{=} n^s\sqrt{6/\pi s(n^2-1)}\ \exp\left\{-6(k-k_0)^2/s(n^2-1)\right\}.$$

For the general case, we shall give here the following illustration Let the weights for a set of successive summations be certain random numbers. For this purpose, let us choose consecutive groups of three numbers from the first basic series (Column 2, Table I, Appendix II). For the first moving summation the weights will be $\alpha_0^{(1)}=5$, $\alpha_1^{(1)}=4$, $\alpha_2^{(1)}=7$; for the second $\alpha_0^{(2)}=3$, $\alpha_1^{(2)}=0$, $\alpha_2^{(2)}=3$, etc. Performing the substitutions indicated by formula (8) we obtain the resulting weights corresponding to A_k of formula (2), $A_k^{(1)}=\alpha_k^{(1)}$, $A_k^{(2)}$, $A_k^{(3)}$, $\cdots$ $A_k^{(10)}$. For each given s, we divide the weights by the largest $A_k^{(s)}$ to obtain the relative weights, $A'_k{}^{(s)}$ (see Table VIII, Appendix II, of the original paper, and Figure 5). The series of quantities $A'_k{}^{(10)}$ does not differ greatly from the Gaussian curve obtained by putting[35]

$$(22) \qquad B'_k{}^{(10)} = 1004 \exp\left\{-\tfrac{1}{2}[(k-9.26)/2.67]^2\right\}.$$

2. The coefficients of correlation between the terms of a coherent series are, in many cases, easy to obtain by using formula (6). For a simple moving summation of n equally weighted items at a time, we have $A_0=A_1=\cdots=A_{n-1}=1$. It is easy to see that

$$(23) \qquad \begin{cases} r_t = (n-|t|)/n, & (|t| \leqq n) \\ r_t = 0, & (|t| \geqq n). \end{cases}$$

From formulas (4), (5), (9) and the properties of $\overset{s}{C}_k$, we find for the s-fold moving summation of two terms, that is, $(n=2)$, that

$$(24) \quad \sigma_y^2 = \sigma_x^2[1 + (\overset{s}{C}_1)^2 + (\overset{s}{C}_2)^2 + \cdots + (\overset{s}{C}_{s-1})^2 + 1] = \sigma_x^2\overset{2s}{C}_s,$$

and

$$(25) \quad E(y_iy_{i+t}) = \sigma_x^2[\overset{s}{C}_0\overset{s}{C}_t + \overset{s}{C}_1\overset{s}{C}_{t+1} + \cdots + \overset{s}{C}_{s-t}\overset{s}{C}_s] = \sigma_x^2\overset{2s}{C}_{s-t}.$$

Hence

$$(26) \qquad r_t = \overset{2s}{C}_{s-t}/\overset{2s}{C}_s = \frac{s(s-1)\cdots(s-t+1)}{(s+1)(s+2)\cdots(s+t)}.$$

Consider another case. Let us form, from a basic series, a coherent series by the scheme:

[35] Let us pass a second degree parabola through $A_0'^{(10)}$, $A_1'^{(10)}$, $A_2'^{(10)}$; another through $A_2'^{(10)}$, $A_3'^{(10)}$, $A_4'^{(10)}$; etc. Denote the area of this figure by S, its maximum ordinate by y_0, and the abscissa bisecting the area by k_0. Then, in the Gaussian equation

$$B'_k{}^{(10)} = [S/\sigma\sqrt{2\pi}]\exp\{-\tfrac{1}{2}[(k-k_0)/\sigma]^2\},$$

all of the parameters are known, since

$$k_0 = 9.26,\ y_0 = S/\sigma\sqrt{2\pi} = 1004,$$

and hence

$$\sigma = S/y_0\sqrt{2\pi} = 2.67$$

$$(27) \qquad y_i = D \sum_{k=0}^{2k_0} x_{i-k} \phi[(k - k_0)/\sigma],$$

where $\phi(t)$ is given by formula (10), D is a constant, and k_0 is a number large enough so that $\phi(t)$ can be neglected for $|t| > k_0/\sigma$. Then, from (6), the coefficient of correlation is

$$(28) \qquad r_t = \frac{\sum_{k=0}^{2k_0-t} \phi[(k - k_0)/\sigma]\phi[(k - k_0 + t)/\sigma]}{\sum_{k=0}^{2k_0} \{\phi[(k - k_0)/\sigma]\}^2}.$$

If σ is sufficiently large we can substitute integrals for the summations in (28). Inserting $z = k - k_0$, we then obtain the approximation formula

$$(29) \qquad r_t = \frac{\int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\frac{z^2 + (z+t)^2}{\sigma^2}\right] dz}{\int_{-\infty}^{+\infty} \exp\left[-\frac{z^2}{\sigma^2}\right] dz}$$

$$= \exp(-t^2/4\sigma^2) = \frac{\phi(t/\sigma\sqrt{2})}{\phi(0)}.$$

Inasmuch as Model III is formed by the scheme of formula (27), with $D = 10^4$, $k_0 = 48$, and $\sigma = 10$, we can calculate the correlation function by formula (29). The values for $[\phi(t/\sqrt{200})]/\phi(0)$, $(t = 0, 1, 2, \cdots)$, were calculated with the aid of Sheppard's tables.[36] The symbol $R_{t(III)}$, instead of $r_{t(III)}$, indicates that an approximate, and not an exact, formula was used in the calculation.

Model IVa was obtained by a 12-fold moving summation of two items; therefore, its correlation function, $r_{t(IVa)}$, is obtained by using formula (26), which, if we consider (11), gives $R_{t(IVa)} = \phi(t/\sqrt{6})/\phi(0)$. The discrepancies between the two results are rather small. The correspondence between $r_{t(IVb)}$ and $R_{t(IVb)}$ is somewhat less, as is also that between $r_{t(IVc)}$ and $R_{t(IVc)}$. Both sets were computed by formula (45) (see next paragraph), but for the calculation of $r_{t(IVb)}$ and $r_{t(IVc)}$ the actual values of $r_{t(IVa)}$ were used as the base, while for $R_{t(IVb)}$ and $R_{t(IVc)}$ the approximate values of $R_{t(IVa)}$, obtained by the Gaussian formula, were used. Even here the discrepancies are not very great when regarded from the same point of view (see Figure 6).

[36] *Tables for Statisticians and Biometricians*, ed. by K. Pearson, Cambridge, 1914, Table II.

Finally, for Model II, corresponding to the scheme

$$y_i = x_i^{(2)} = \sum_{k=0}^{18} A_k^{(2)} x_{i+k} + 50, \tag{30}$$

$$(A_k^{(2)} = 1, 2, \cdots, 9, 10, 9, \cdots, 2, 1),$$

the coefficients of correlation can be obtained directly from formula (6). It is worth noting that, even in this case, a good approximation, $R_{t(II)}$, can be obtained by the use of the Gaussian curve, the equation being

$$R_{t(II)} = \exp\left[-\tfrac{1}{2}(t/5.954)^2\right], \tag{31}$$

where $\sigma = 5.954$ was obtained by equating the areas of the Gaussian curve and of the empirical curve, and the computations were carried through with the help of Simpson's rule (see Figure 6).

A few more words may be said about the correlational function for the weights, $A'^{(10)}_k$, of our example of the crossing of random weights (see Section 2 of the text, Appendix, Section 1, and Figure 6). The exact values of the coefficients of correlation (r_t) were found by formula (6), while approximate values (R_t) were obtained from the equation

$$R_t = \exp\left[-\tfrac{1}{2}(t/3.727)^2\right], \tag{32}$$

which was obtained in the same manner as was equation (31). Both the exact and approximate values are given in Table 1 (see also Figure 6). Also, let us note that, from the equation of the Gaussian curve which approximates the weights, $A'^{(10)}_k$ (see formula (22) above), it is possible to find an approximate expression for the coefficients of correlation by using formula (29). An expression analogous to (32) would be obtained, but instead of $\sigma = 3.727$, we would have $\sigma = 3.776$. The correlation coefficients are only slightly less accurate than those found from formula (32), the deviations are all of one sign, and none is greater than 0.009.

Let us make one more observation. If a chance variable $y_i = u_i + v_i$, where u_i is a coherent series and v_i is a random series, it is easy to show that

$$E(y_i y_{i+t}) = E(u_i u_{i+t}), \tag{33}$$

$$(Ey_i^2) = \sigma_u^2 + \sigma_v^2, \tag{34}$$

$$r_{y_i,y_{i+t}} = \frac{r_{u_i,u_{i+t}}}{1 + (\sigma_v^2/\sigma_u^2)}, \tag{35}$$

where $E(u_i)$ and $E(v_i)$ are taken equal to zero.

If $r_{u_i,u_{i+t}}$ lies along the Gaussian curve, then $r_{y_i,y_{i+t}}$ will lie along

a similar curve with ordinates proportionally reduced, except that r_0 will, as formerly, equal unity; the *chapeau de gendarme* has taken on the spike of the *budenovka* (a Soviet military cap). It is to be expected that this figure and the analogous figures for the correlation function of the differences (formula (45) of the following paragraph) will be encountered in the investigation of empirical series.[37]

3. Let us now investigate the differences of various orders of the series y_i, i.e., $\Delta^\alpha y_i$, $\Delta^\beta y_i$, and their coefficients of correlation.[38] As before, let

$$E(y_i) = 0;\ E(y_i^2) = \sigma_y^2 = \text{constant};\ E(y_i y_{i+t})/\sigma_y^2 = r_t, \tag{36}$$

where r_t is supposed to be independent of i. Let us introduce the notation

$$r_t^{(\alpha,\beta)} = r_{\Delta^\alpha y_i \Delta^\beta y_{i+t}}, \qquad (\alpha \geqq \beta); \tag{37}$$

and, in particular,

$$r_t^{(\alpha,\alpha)} = r_{\Delta^\alpha y_i \Delta^\alpha y_{i+t}} \qquad r_t^{(\alpha,0)} = r_{\Delta^\alpha y_i, y_{i+t}}. \tag{38}$$

By using the equality

$$C_k^{2\alpha} = C_{\alpha-k}^{\alpha} C_0^{\alpha} + C_{\alpha-(k-1)}^{\alpha} C_1^{\alpha} + \cdots + C_{\alpha-1}^{\alpha} C_{k-1}^{\alpha} + C_\alpha^\alpha C_k^\alpha, \tag{39}$$

it can be shown that

$$\begin{aligned} \sigma_{\Delta^\alpha y_i}^2 &= E[(\Delta^\alpha y_i)^2] = E[(C_\alpha^\alpha y_{i+\alpha} - C_{\alpha-1}^\alpha y_{i+\alpha-1} + C_{\alpha-2}^\alpha y_{i+\alpha-2} \\ &\quad - \cdots + (-1)^\alpha C_0^\alpha y_i)^2] = (-1)^\alpha \Delta^{2\alpha} r_{-\alpha} \sigma_y^2. \end{aligned} \tag{40}$$

From (40), by using the equality,

$$C_k^{\alpha+\beta} = C_{\alpha-k}^{\alpha} C_0^{\beta} + C_{\alpha-k+1}^{\alpha} C_1^{\beta} + \cdots + C_{\alpha-1}^{\alpha} C_{k-1}^{\beta} + C_k^{\beta}, \tag{41}$$

we obtain

$$E[\Delta^\alpha y_i \Delta^\beta y_{i+t}] = (-)^\alpha \Delta^{\alpha+\beta} r_{t-\alpha} \sigma_y^2, \qquad (\alpha \geqq \beta), \tag{42}$$

and from this we have

$$r_t^{(\alpha,\beta)} = \frac{(-1)^\alpha \Delta^{\alpha+\beta} r_{t-\alpha}}{\sqrt{(-1)^{\alpha+\beta} \Delta^{2\alpha} r_{-\alpha} \Delta^{2\beta} r_{-\beta}}}. \tag{43}$$

In the same manner we obtain

$$r_t^{(\alpha,0)} = \frac{(-1)^\alpha \Delta^\alpha r_{t-\alpha}}{\sqrt{(-1)^\alpha \Delta^{2\alpha} r_{-\alpha}}}, \tag{44}$$

[37] Cf. Figure 19 of Yule, "Why Do We Sometimes Get Nonsense-Correlations · · · ," *loc. cit.*, p. 43.

[38] Cf. O. Anderson, "Über ein neues Verfahren bei Anwendung der 'Variate-Difference' Methode," *Biometrika*, Vol. 15, 1923, pp. 142 ff.

and, as a special case of formula (43), we have

$$r_t^{(\alpha,\alpha)} = \frac{\Delta^{2\alpha} r_{t-\alpha}}{\Delta^{2\alpha} r_{-\alpha}}. \tag{45}$$

By this formula and by (26) the correlation coefficients for Models IV*b* and IV*c* have been computed.

4. Let us now prove Theorem A (see Section VII above). The regression coefficient of $\Delta^2 y_i$ on y_{i+1} being

$$E(\Delta^2 y_i \cdot y_{i+1})/\sigma_y^2 = -2(1 - r_1),$$

we obtain the approximate equation

$$\Delta^2 y_i \overset{a}{=} -2(1 - r_1) y_{i+1}.$$

Whence

$$y_{i+2} \overset{a}{=} 2r_1 y_{i+1} - y_i, \tag{46}$$

the errors of both equations being evidently identical. If we denote this error by α_2 and put $\beta_2 = \alpha_2/\sigma_y$, we may apply the well-known formula of correlation theory and thus obtain

$$E(\beta_2^2) = E(\Delta^2 y_i)(1 - \rho_1^2)/\sigma_y^2 = (1 - \rho_1^2)(6 - 8r_1 + 2r_2). \tag{47}$$

Now, under the suppositions of Theorem A,

$$\lim_{n\to\infty} \beta_2^2 = 0;$$

whence, applying Tchebycheff's theorem, we see that β_2 has the stochastical limit $E(\beta_2) = 0$, $(n \to \infty)$; that is, ϵ and η being arbitrarily small, the probability

$$P\{\,|\beta_2| > \epsilon\} < \eta,$$

provided n is sufficiently large.

On the other hand, if we put

$$\begin{aligned} y_{i+2} &= 2r_1 y_{i+1} - y_i + \alpha_2, \\ y_{i+3} &= 2r_1 y_{i+2} - y_{i+1} + \alpha_3, \\ &\ldots\ldots\ldots\ldots\ldots \\ y_{i+s} &= 2r_1 y_{i+s-1} - y_{i+s-2} + \alpha_s; \end{aligned}$$

and if we insert y_{i+2} in the second equation of this system, y_{i+3} in the third, and so on, we obtain, after reduction,

$$y_{i+s} = C_1 y_{i+1} + C_2 y_i + \lambda_s \sigma_y, \tag{48}$$

where

$$\lambda_s = a_0\beta_s + a_1\beta_{s-1} + \cdots + a_{s-2}\beta_2, \tag{49}$$

the values $a_0, a_1, \cdots a_{s+2}$, being determined by the conditions

$$\begin{cases} a_0 = 1, \quad a_1 = 2r_1, \\ a_{k+2} = 2r_1 a_{k+1} - a_k. \end{cases} \tag{50}$$

This equation is identical with

$$y_{i+2} = 2r_1 y_{i+1} - y_i, \tag{51}$$

that is, with (46) considered as a precise equation. The solution of (51) or (50) can be obtained easily. We find

$$y_t = A \cos (2\pi t/L) + B \sin (2\pi t/L), \tag{52}$$

and

$$a_k = C \cos (2\pi k/L) + D \sin (2\pi k/L), \tag{53}$$

where we have put

$$\cos (2\pi/L) = r_1, \tag{54}$$

L being the period of the respective sinusoid. It is evident that, under the assumptions of Theorem A, ($|r_1| \leqq \lambda < 1$), we shall have

$$L \leqq 2\pi/\text{arc} \cos \lambda = H = \text{constant}. \tag{55}$$

Two sinusoids must now be considered. The first, which will be denoted by S_1, is determined by (51), or (52), and the initial points y_i, y_{i+1}. It is evident that $C_1 y_{i+1} + C_2 y_i$ in (48) is the ordinate of S_1 which could be obtained in this form from (51) by successive substitutions. The deviation of the actual value of y_{i+s} from S_1 is $\lambda_s \sigma_y$ as given by (48) and (49), the coefficients a_k being the ordinates of the second sinusoid (S_2) determined by (50) or (53), and the initial values $a_0 = 1$, $a_1 = 2r_1$. But, if we put, in (53), $k = 0$, and then $k = 1$, we obtain

$$C = 1, \quad D = r_1/\sqrt{1 - r_1^2}.$$

Hence the amplitude of S_2 is

$$\sqrt{C^2 + D^2} = 1/\sqrt{1 - r_1^2} \leqq 1/\sqrt{1 - \lambda^2} = K = \text{constant}.$$

$$C = 1, \quad D = r_1/\sqrt{1 - r_1^2}.$$

Thus, taking into account that every a_k in (49) has an upper limit $\leqq \sqrt{C^2+D^2}$, $(n\to\infty)$ and remembering the theorems of my *Metron* memoir[39] we conclude that λ_s has the stochastical limit $=0$ and that, ϵ and η being arbitrarily small and s arbitrarily large, the probability that the conditions

$$\lambda_2 < \epsilon,\ \lambda_3 < \epsilon, \cdots, \lambda_s < \epsilon,$$

are simultaneously satisfied will be $>1-\eta$ provided n is sufficiently large. The formulas (54) and (55) complete the proof.

To prove theorem B, we may proceed here as follows: It is seen by (43) and (45) that the correlation coefficient between y_i and y_{i+1} is

$$r_1 = r_1^{(m,m)} = \frac{\Delta^{2m}r_{-m+1}}{\Delta^{2m}r_{-m}}, \tag{56}$$

and the correlation coefficient between $\Delta^2 y_i$ and y_{i+1} is given by

$$\rho_1 = r_1^{(m+2,m)} = \frac{(-1)^m\Delta^{2m+m}r_{-(m+1)}}{\sqrt{\Delta^{2(m+2)}r_{-(m+2)}\Delta^{2m}r_{-m}}}, \tag{57}$$

where, using (26) we must put

$$r_t = \frac{C_{n-t}^{2n}}{C_n^{2n}}.$$

On the other hand, r_{-t} being equal to r_t, it can easily be seen that

$$C_n^{2n}\Delta^{2m}r_{-m} = \sum_{k=0}^{2m}(-1)^k C_k^{2m} C_{n-m+k}^{2n} = A_{n+m}, \tag{58}$$

where A_{n+m} is the coefficient of x^{n+m} in the expansion of $(1+x)^{2n}(1-x)^{2m}$. Applying Cauchy's theorem we have

$$A_{n+m}\frac{1}{2\pi i}\int_{|1|}\frac{(1+x)^{2n}(1-x)^{2m}}{x^{n+m+1}}\,dx.$$

If we put $x=e^{i\phi}$, we obtain, after reduction,

$$A_{n+m} = \left[(-1)^m 2^{2(n+m)}/\pi\right]\int_0^{\pi}\cos^{2n}\phi\,\sin^{2m}\phi\,d\phi.$$

[39] E. Slutzky, "Über Stochastische Asymptoten und Grenzwerte," *Metron*, Vol. 5, N. 3, 1925, pp. 61–64.

Hence

$$(59)\quad A_{n+m} = \frac{(-1)^m 2^{n+m} 1\cdot 3 \cdots (2n-1)\cdot 1\cdot 3 \cdots (2m-1)}{1\cdot 2\cdot 3 \cdots (n+m)}.$$

Thus we obtain, by (58),

$$(60)\qquad \Delta^{2m} r_{-m} = \frac{(-1)^m 2^m 1\cdot 3 \cdots (2m-1)}{(n+1)(n+2)\cdots(n+m)}.$$

If we notice that

$$\Delta^{2m+2} r_{-(m+1)} = \Delta^{2m} r_{-(m+1)} - 2\Delta^{2m} r_{-m} + \Delta^{2m} r_{-(m-1)},$$

where, evidently,

$$\Delta^{2m} r_{-(m+1)} = \Delta^{2m} r_{-(m-1)},$$

we get

$$\begin{aligned} \Delta^{2m} r_{-m+1} &= \tfrac{1}{2}\Delta^{2(m+1)} r_{-(m+1)} + \Delta^{2m} r_{-m} \\ (61)\qquad &= \frac{(-1)^m 2^m 1\cdot 3 \cdots (2m-1)(n-m)}{(n+1)(n+2)\cdots(n+m+1)}, \end{aligned}$$

so that, by (56), (57), (60), and (61), we obtain

$$(62)\qquad r_1 = (n-m)/(n+m+1)$$

and[40]

$$(63)\qquad \rho_1 = -\sqrt{\frac{(2m+1)(n+m+2)}{(2m+3)(n+m+1)}}.$$

Now, it is evident that, n/m being constant,

$$r_1 < \frac{n/m - 1}{n/m + 1} < 1$$

and

$$\rho_1 \rightarrow -1, \qquad (n \rightarrow \infty),$$

which proves Theorem B.

CORRECTIONS OF BASIC DATA

The tables of figures which contain the series used in the present investigation are to be found in the original paper (*loc. cit.*, pp. 57–64). As the preparation of them has involved a great deal of time and labor

[40] To apply this formula in the case of No. V, we should notice that Model III is approximately equivalent to the series (y_1) of Theorem B, with $n = 400$.

and as it may be expected that someone will make use of them for the purpose of analogous studies, we give here correct readings for the *errata* found after the figures had been published. (Those relating to Table VI are immaterial and are omitted here.)

Table	Column					
	1	2	6	7	8	10
I	300				−453	
	418			−361		
	637				332	
	638			− 10	461	
	639		1367	451	295	
	807			255		
	819				211	
	820			496	−252	
	821		2290	244	−488	
	971				− 99	
	972			263	− 98	
	973		2134	165	−186	
III	23	− 4551				
	72	−20219				
IX	9		0.0007			
	11					0.0059
	13					0.0001
	14					0.0000

EUGEN SLUTZKY

Mathematical Institute of the Moscow State University
Moscow, U.S.S.R.

MR. KEYNES AND THE "CLASSICS"; A SUGGESTED INTERPRETATION[1]

By J. R. Hicks

I

It will be admitted by the least charitable reader that the entertainment value of Mr. Keynes' *General Theory of Employment* is considerably enhanced by its satiric aspect. But it is also clear that many readers have been left very bewildered by this Dunciad. Even if they are convinced by Mr. Keynes' arguments and humbly acknowledge themselves to have been "classical economists" in the past, they find it hard to remember that they believed in their unregenerate days the things Mr. Keynes says they believed. And there are no doubt others who find their historic doubts a stumbling block, which prevents them from getting as much illumination from the positive theory as they might otherwise have got.

One of the main reasons for this situation is undoubtedly to be found in the fact that Mr. Keynes takes as typical of "Classical economics" the later writings of Professor Pigou, particularly *The Theory of Unemployment.* Now *The Theory of Unemployment* is a fairly new book, and an exceedingly difficult book; so that it is safe to say that it has not yet made much impression on the ordinary teaching of economics. To most people its doctrines seem quite as strange and novel as the doctrines of Mr. Keynes himself; so that to be told that he has believed these things himself leaves the ordinary economist quite bewildered.

For example, Professor Pigou's theory runs, to a quite amazing extent, in real terms. Not only is his theory a theory of real wages and unemployment; but numbers of problems which anyone else would have preferred to investigate in money terms are investigated by Professor Pigou in terms of "wage-goods." The ordinary classical economist has no part in this *tour de force.*

But if, on behalf of the ordinary classical economist, we declare that he would have preferred to investigate many of those problems in money terms, Mr. Keynes will reply that there is no classical theory of money wages and employment. It is quite true that such a theory cannot easily be found in the textbooks. But this is only because most of the textbooks were written at a time when general changes in money wages in a closed system did not present an important problem. There can be little doubt that most economists have thought that they had

[1] Based on a paper which was read at the Oxford meeting of the Econometric Society (September, 1936) and which called forth an interesting discussion. It has been modified subsequently, partly in the light of that discussion, and partly as a result of further discussion in Cambridge.

a pretty fair idea of what the relation between money wages and employment actually was.

In these circumstances, it seems worth while to try to construct a typical "classical" theory, built on an earlier and cruder model than Professor Pigou's. If we can construct such a theory, and show that it does give results which have in fact been commonly taken for granted, but which do not agree with Mr. Keynes' conclusions, then we shall at last have a satisfactory basis of comparison. We may hope to be able to isolate Mr. Keynes' innovations, and so to discover what are the real issues in dispute.

Since our purpose is comparison, I shall try to set out my typical classical theory in a form similar to that in which Mr. Keynes sets out his own theory; and I shall leave out of account all secondary complications which do not bear closely upon this special question in hand. Thus I assume that I am dealing with a short period in which the quantity of physical equipment of all kinds available can be taken as fixed. I assume homogeneous labour. I assume further that depreciation can be neglected, so that the output of investment goods corresponds to new investment. This is a dangerous simplification, but the important issues raised by Mr. Keynes in his chapter on user cost are irrelevant for our purposes.

Let us begin by assuming that w, the rate of money wages per head, can be taken as given.

Let x, y, be the outputs of investment goods and consumption goods respectively, and N_x, N_y, be the numbers of men employed in producing them. Since the amount of physical equipment specialised to each industry is given, $x = f_x(N_x)$ and $y = f_y(N_y)$, where f_x, f_y, are *given* functions.

Let M be the *given* quantity of money.

It is desired to determine N_x and N_y.

First, the price-level of investment goods = their marginal cost = $w(dN_x/dx)$. And the price-level of consumption goods = their marginal cost = $w(dN_y/dy)$.

Income earned in investment trades (value of investment, or simply Investment) = $wx(dN_x/dx)$. Call this I_x.

Income earned in consumption trades = $wy(dN_y/dy)$.

Total Income = $wx(dN_x/dx) + wy(dN_y/dy)$. Call this I.

I_x is therefore a given function of N_x, I of N_x and N_y. Once I and I_x are determined, N_x and N_y can be determined.

Now let us assume the "Cambridge Quantity equation"—that there is some definite relation between Income and the demand for money. Then, approximately, and apart from the fact that the demand for money may depend not only upon total Income, but also upon its dis-

tribution between people with relatively large and relatively small demands for balances, we can write

$$M = kI.$$

As soon as k is given, total Income is therefore determined.

In order to determine I_x, we need two equations. One tells us that the amount of investment (looked at as demand for capital) depends upon the rate of interest:

$$I_x = C(i).$$

This is what becomes the marginal-efficiency-of-capital schedule in Mr. Keynes' work.

Further, Investment = Saving. And saving depends upon the rate of interest and, if you like, Income. $\therefore I_x = S(i, I)$. (Since, however, Income is already determined, we do not need to bother about inserting Income here unless we choose.)

Taking them as a system, however, we have three fundamental equations,

$$M = kI, \quad I_x = C(i), \quad I_x = S(i, I),$$

to determine three unknowns, I, I_x, i. As we have found earlier, N_x and N_y can be determined from I and I_x. Total employment, $N_x + N_y$, is therefore determined.

Let us consider some properties of this system. It follows directly from the first equation that as soon as k and M are given, I is completely determined; that is to say, total income depends directly upon the quantity of money. Total employment, however, is not necessarily determined at once from income, since it will usually depend to some extent upon the proportion of income saved, and thus upon the way production is divided between investment and consumption-goods trades. (If it so happened that the elasticities of supply were the same in each of these trades, then a shifting of demand between them would produce compensating movements in N_x and N_y, and consequently no change in total employment.)

An increase in the inducement to invest (i.e., a rightward movement of the schedule of the marginal efficiency of capital, which we have written as $C(i)$) will tend to raise the rate of interest, and so to affect saving. If the amount of saving rises, the amount of investment will rise too; labour will be employed more in the investment trades, less in the consumption trades; this will increase total employment if the elasticity of supply in the investment trades is greater than that in the consumption-goods trades—diminish it if *vice versa.*

An increase in the supply of money will necessarily raise total income, for people will increase their spending and lending until incomes have risen sufficiently to restore k to its former level. The rise in income

will tend to increase employment, both in making consumption goods and in making investment goods. The total effect on employment depends upon the ratio between the expansions of these industries; and that depends upon the proportion of their increased incomes which people desire to save, which also governs the rate of interest.

So far we have assumed the rate of money wages to be given; but so long as we assume that k is independent of the level of wages, there is no difficulty about this problem either. A rise in the rate of money wages will necessarily diminish employment and raise real wages. For an unchanged money income cannot continue to buy an unchanged quantity of goods at a higher price-level; and, unless the price-level rises, the prices of goods will not cover their marginal costs. There must therefore be a fall in employment; as employment falls, marginal costs in terms of labour will diminish and therefore real wages rise. (Since a change in money wages is always accompanied by a change in real wages in the same direction, if not in the same proportion, no harm will be done, and some advantage will perhaps be secured, if one prefers to work in terms of real wages. Naturally most "classical economists" have taken this line.)

I think it will be agreed that we have here a quite reasonably consistent theory, and a theory which is also consistent with the pronouncements of a recognizable group of economists. Admittedly it follows from this theory that you may be able to increase employment by direct inflation; but whether or not you decide to favour that policy still depends upon your judgment about the probable reaction on wages, and also—in a national area—upon your views about the international standard.

Historically, this theory descends from Ricardo, though it is not actually Ricardian; it is probably more or less the theory that was held by Marshall. But with Marshall it was already beginning to be qualified in important ways; his successors have qualified it still further. What Mr. Keynes has done is to lay enormous emphasis on the qualifications, so that they almost blot out the original theory. Let us follow out this process of development.

II

When a theory like the "classical" theory we have just described is applied to the analysis of industrial fluctuations, it gets into difficulties in several ways. It is evident that total money income experiences great variations in the course of a trade cycle, and the classical theory can only explain these by variations in M or in k, or, as a third and last alternative, by changes in distribution.

(1) Variation in M is simplest and most obvious, and has been relied

on to a large extent. But the variations in M that are traceable during a trade cycle are variations that take place through the banks—they are variations in bank loans; if we are to rely on them it is urgently necessary for us to explain the connection between the supply of bank money and the rate of interest. This can be done roughly by thinking of banks as persons who are strongly inclined to pass on money by lending rather than spending it. Their action therefore tends at first to lower interest rates, and only afterwards, when the money passes into the hands of spenders, to raise prices and incomes. "The new currency, or the increase of currency, goes, not to private persons, but to the banking centers; and therefore, it increases the willingness of lenders to lend in the first instance, and lowers the rate of discount. But it afterwards raises prices; and therefore it tends to increase discount."[2] This is superficially satisfactory; but if we endeavoured to give a more precise account of this process we should soon get into difficulties. What determines the amount of money needed to produce a given fall in the rate of interest? What determines the length of time for which the low rate will last? These are not easy questions to answer.

(2) In so far as we rely upon changes in k, we can also do well enough up to a point. Changes in k can be related to changes in confidence, and it is realistic to hold that the rising prices of a boom occur because optimism encourages a reduction in balances; the fa ling prices of a slump because pessimism and uncertainty dictate an increase. But as soon as we take this step it becomes natural to ask whether k has not abdicated its status as an independent variable, and has not become liable to be influenced by others among the variables in our fundamental equations.

(3) This last consideration is powerfully supported by another, of more purely theoretical character. On grounds of pure value theory, it is evident that the direct sacrifice made by a person who holds a stock of money is a sacrifice of interest; and it is hard to believe that the marginal principle does not operate at all in this field. As Lavington put it: "The quantity of resources which (an individual) holds in the form of money will be such that the unit of money which is just and only just worth while holding in this form yields him a return of convenience and security equal to the yield of satisfaction derived from the marginal unit spent on consumables, and equal also to the net rate of interest."[3] The demand for money depends upon the rate of interest! The stage is set for Mr. Keynes.

[2] Marshall, *Money, Credit, and Commerce*, p. 257.

[3] Lavington, *English Capital Market*, 1921, p. 30. See also Pigou, "The Exchange-value of Legal-tender Money," in *Essays in Applied Economics*, 1922, pp. 179–181.

As against the three equations of the classical theory,

$$M = kI, \quad I_x = C(i), \quad I_x = S(i, I),$$

Mr. Keynes begins with three equations,

$$M = L(i), \quad I_x = C(i), \quad I_x = S(I).$$

These differ from the classical equations in two ways. On the one hand, the demand for money is conceived as depending upon the rate of interest (Liquidity Preference). On the other hand, any possible influence of the rate of interest on the amount saved out of a given income is neglected. Although it means that the third equation becomes the multiplier equation, which performs such queer tricks, nevertheless this second amendment is a mere simplification, and ultimately insignificant.[4] It is the liquidity preference doctrine which is vital.

For it is now the rate of interest, not income, which is determined by the quantity of money. The rate of interest set against the schedule of the marginal efficiency of capital determines the value of investment; that determines income by the multiplier. Then the volume of employment (at given wage-rates) is determined by the value of investment and of income which is not saved but spent upon consumption goods.

It is this system of equations which yields the startling conclusion, that an increase in the inducement to invest, or in the propensity to consume, will not tend to raise the rate of interest, but only to increase employment. In spite of this, however, and in spite of the fact that quite a large part of the argument runs in terms of this system, and this system alone, *it is not the General Theory*. We may call it, if we like, Mr. Keynes' *special theory*. The General Theory is something appreciably more orthodox.

Like Lavington and Professor Pigou, Mr. Keynes does not in the end believe that the demand for money can be determined by one variable alone—not even the rate of interest. He lays more stress on it than they did, but neither for him nor for them can it be the only variable to be considered. The dependence of the demand for money on interest does not, in the end, do more than qualify the old de-

[4] This can be readily seen if we consider the equations

$$M = kI, \quad I_x = C(i), \quad I_x = S(I),$$

which embody Mr. Keynes' second amendment without his first. The third equation is already the multiplier equation, but the multiplier is shorn of his wings. For since I still depends only on M, I_x now depends only on M, and it is impossible to increase investment without increasing the willingness to save or the quantity of money. The system thus generated is therefore identical with that which, a few years ago, used to be called the "Treasury View." But Liquidity Preference transports us from the "Treasury View" to the "General Theory of Employment."

pendence on income. However much stress we lay upon the "speculative motive," the "transactions" motive must always come in as well.

Consequently we have for the General Theory

$$M = L(I, i), \quad I_x = C(i), \quad I_x = S(I).$$

With this revision, Mr. Keynes takes a big step back to Marshallian orthodoxy, and his theory becomes hard to distinguish from the revised and qualified Marshallian theories, which, as we have seen, are not new. Is there really any difference between them, or is the whole thing a sham fight? Let us have recourse to a diagram (Figure 1).

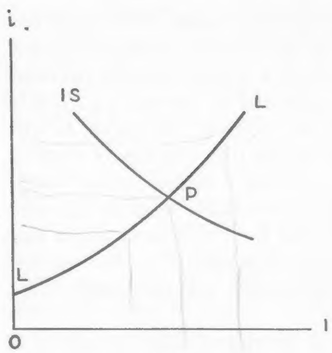


FIGURE 1

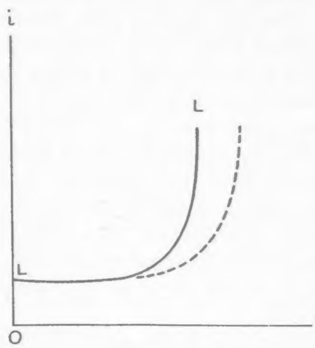


FIGURE 2

Against a given quantity of money, the first equation, $M = L(I, i)$, gives us a relation between Income (I) and the rate of interest (i). This can be drawn out as a curve (LL) which will slope upwards, since an increase in income tends to raise the demand for money, and an increase in the rate of interest tends to lower it. Further, the second two equations taken together give us another relation between Income and interest. (The marginal-efficiency-of-capital schedule determines the value of investment at any given rate of interest, and the multiplier tells us what level of income will be necessary to make savings equal to that value of investment.) The curve IS can therefore be drawn showing the relation between Income and interest which must be maintained in order to make saving equal to investment.

Income and the rate of interest are now determined together at P, the point of intersection of the curves LL and IS. They are determined together; just as price and output are determined together in the modern theory of demand and supply. Indeed, Mr. Keynes' innovation is closely parallel, in this respect, to the innovation of the marginalists.

The quantity theory tries to determine income without interest, just as the labour theory of value tried to determine price without output; each has to give place to a theory recognising a higher degree of interdependence.

III

But if this is the real "General Theory," how does Mr. Keynes come to make his remarks about an increase in the inducement to invest not raising the rate of interest? It would appear from our diagram that a rise in the marginal-efficiency-of-capital schedule must raise the curve *IS*; and, therefore, although it will raise Income and employment, it will also raise the rate of interest.

This brings us to what, from many points of view, is the most important thing in Mr. Keynes' book. It is not only possible to show that a given supply of money determines a certain relation between Income and interest (which we have expressed by the curve *LL*); it is also possible to say something about the shape of the curve. It will probably tend to be nearly horizontal on the left, and nearly vertical on the right. This is because there is (1) some minimum below which the rate of interest is unlikely to go, and (though Mr. Keynes does not stress this) there is (2) a maximum to the level of income which can possibly be financed with a given amount of money. If we like we can think of the curve as approaching these limits asymptotically (Figure 2).

Therefore, if the curve *IS* lies well to the right (either because of a strong inducement to invest or a strong propensity to consume), *P* will lie upon that part of the curve which is decidedly upward sloping, and the classical theory will be a good approximation, needing no more than the qualification which it has in fact received at the hands of the later Marshallians. An increase in the inducement to invest will raise the rate of interest, as in the classical theory, but it will also have some subsidiary effect in raising income, and therefore employment as well. (Mr. Keynes in 1936 is not the first Cambridge economist to have a temperate faith in Public Works.) But if the point *P* lies to the left of the *LL* curve, then the *special* form of Mr. Keynes' theory becomes valid. A rise in the schedule of the marginal efficiency of capital only increases employment, and does not raise the rate of interest at all. We are completely out of touch with the classical world.

The demonstration of this minimum is thus of central importance. It is so important that I shall venture to paraphrase the proof, setting it out in a rather different way from that adopted by Mr. Keynes.[5]

If the costs of holding money can be neglected, it will always be

[5] Keynes, *General Theory*, pp. 201–202.

profitable to hold money rather than lend it out, if the rate of interest is not greater than zero. Consequently the rate of interest must always be positive. In an extreme case, the shortest short-term rate may perhaps be nearly zero. But if so, the long-term rate must lie above it, for the long rate has to allow for the risk that the short rate may rise during the currency of the loan, and it should be observed that the short rate can only rise, it cannot fall.[6] This does not only mean that the long rate must be a sort of average of the probable short rates over its duration, and that this average must lie above the current short rate. There is also the more important risk to be considered, that the lender on long term may desire to have cash before the agreed date of repayment, and then, if the short rate has risen meanwhile, he may be involved in a substantial capital loss. It is this last risk which provides Mr. Keynes' "speculative motive" and which ensures that the rate for loans of indefinite duration (which he always has in mind as *the* rate of interest) cannot fall very near zero.[7]

It should be observed that this minimum to the rate of interest applies not only to one curve *LL* (drawn to correspond to a particular quantity of money) but to any such curve. If the supply of money is increased, the curve *LL* moves to the right (as the dotted curve in Figure 2), but the horizontal parts of the curve are almost the same. Therefore, again, it is this doldrum to the left of the diagram which upsets the classical theory. If *IS* lies to the right, then we can indeed increase employment by increasing the quantity of money; but if *IS* lies to the left, we cannot do so; merely monetary means will not force down the rate of interest any further.

So the General Theory of Employment is the Economics of Depression.

[6] It is just conceivable that people might become so used to the idea of very low short rates that they would not be much impressed by this risk; but it is very unlikely. For the short rate may rise, either because trade improves, and income expands; or because trade gets worse, and the desire for liquidity increases. I doubt whether a monetary system so elastic as to rule out both of these possibilities is really thinkable.

[7] Nevertheless something more than the "speculative motive" is needed to account for the system of interest rates. The shortest of all short rates must equal the relative valuation, at the margin, of money and such a bill; and the bill stands at a discount mainly because of the "convenience and security" of holding money—the inconvenience which may possibly be caused by not having cash immediately available. It is the chance that you may want to discount the bill which matters, not the chance that you will then have to discount it on unfavourable terms. The "precautionary motive," not the "speculative motive," is here dominant. But the prospective terms of rediscounting are vital, when it comes to the *difference* between short and long rates.

IV

In order to elucidate the relation between Mr. Keynes and the "Classics," we have invented a little apparatus. It does not appear that we have exhausted the uses of that apparatus, so let us conclude by giving it a little run on its own.

With that apparatus at our disposal, we are no longer obliged to make certain simplifications which Mr. Keynes makes in his exposition. We can reinsert the missing i in the third equation, and allow for any possible effect of the rate of interest upon saving; and, what is much more important, we can call in question the sole dependence of investment upon the rate of interest, which looks rather suspicious in the second equation. Mathematical elegance would suggest that we ought to have I and i in all three equations, if the theory is to be really General. Why not have them there like this:

$$M = L(I, i), \quad I_x = C(I, i), \quad I_x = S(I, i)?$$

Once we raise the question of Income in the second equation, it is clear that it has a very good claim to be inserted. Mr. Keynes is in fact only enabled to leave it out at all plausibly by his device of measuring everything in "wage-units," which means that he allows for changes in the marginal-efficiency-of-capital schedule when there is a change in the level of money wages, but that other changes in Income are deemed not to affect the curve, or at least not in the same immediate manner. But why draw this distinction? Surely there is every reason to suppose that an increase in the demand for consumers' goods, arising from an increase in employment, will often directly stimulate an increase in investment, at least as soon as an expectation develops that the increased demand will continue. If this is so, we ought to include I in the second equation, though it must be confessed that the effect of I on the marginal efficiency of capital will be fitful and irregular.

The Generalized General Theory can then be set out in this way. Assume first of all a given total money Income. Draw a curve CC showing the marginal efficiency of capital (in money terms) at that given Income; a curve SS showing the supply curve of saving at that *given* Income (Figure 3). Their intersection will determine the rate of interest which makes savings equal to investment at that level of income. This we may call the "investment rate."

If Income rises, the curve SS will move to the right; probably CC will move to the right too. If SS moves more than CC, the investment rate of interest will fall; if CC more than SS, it will rise. (How much it rises and falls, however, depends upon the elasticities of the CC and SS curves.)

The IS curve (drawn on a separate diagram) now shows the relation

between Income and the corresponding investment rate of interest. It has to be confronted (as in our earlier constructions) with an *LL* curve showing the relation between Income and the "money" rate of interest; only we can now generalise our *LL* curve a little. Instead of assuming, as before, that the supply of money is given, we can assume that there is a given monetary system—that up to a point, but only up to a point, monetary authorities will prefer to create new money rather than allow interest rates to rise. Such a generalised *LL* curve will then slope upwards only gradually—the elasticity of the curve depending on the elasticity of the monetary system (in the ordinary monetary sense).

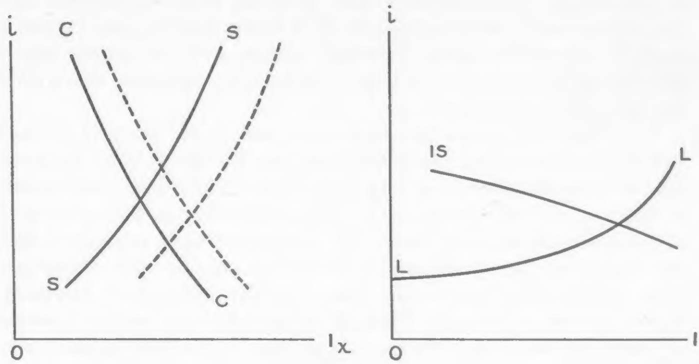


FIGURE 3

As before, Income and interest are determined where the *IS* and *LL* curves intersect—where the investment rate of interest equals the money rate. Any change in the inducement to invest or the propensity to consume will shift the *IS* curve; any change in liquidity preference or monetary policy will shift the *LL* curve. If, as the result of such a change, the investment rate is raised above the money rate, Income will tend to rise; in the opposite case, Income will tend to fall; the extent to which Income rises or falls depends on the elasticities of the curves.[8]

[8] Since $C(I, i) = S(I, i)$, $$\frac{dI}{di} = -\frac{\partial S/\partial i - \partial C/\partial i}{\partial S/\partial I - \partial C/\partial I}.$$

The savings investment market will not be stable unless $\partial S/\partial i + (-\partial C/\partial i)$ is positive. I think we may assume that this condition is fulfilled.

If $\partial S/\partial i$ is positive, $\partial C/\partial i$ negative, $\partial S/\partial I$ and $\partial C/\partial I$ positive (the most probable state of affairs), we can say that the *IS* curve will be more elastic, the

When generalised in this way, Mr. Keynes' theory begins to look very like Wicksell's; this is of course hardly surprising.[9] There is indeed one special case where it fits Wicksell's construction absolutely. If there is "full employment" in the sense that any rise in Income immediately calls forth a rise in money wage rates; then it is *possible* that the CC and SS curves may be moved to the right to exactly the same extent, so that IS is horizontal. (I say possible, because it is not unlikely, in fact, that the rise in the wage level may create a presumption that wages will rise again later on; if so, CC will probably be shifted more than SS, so that IS will be upward sloping.) However that may be, if IS is horizontal, we do have a perfectly Wicksellian construction;[10] the investment rate becomes Wicksell's *natural rate*, for in this case it may be thought of as determined by real causes; if there is a perfectly elastic monetary system, and the money rate is fixed below the natural rate, there is cumulative inflation; cumulative deflation if it is fixed above.

This, however, is now seen to be only one special case; we can use our construction to harbour much wider possibilities. If there is a great deal of unemployment, it is very likely that $\partial C/\partial I$ will be quite small; in that case IS can be relied upon to slope downwards. This is the sort of Slump Economics with which Mr. Keynes is largely concerned. But one cannot escape the impression that there may be other conditions when expectations are tinder, when a slight inflationary tendency lights them up very easily. Then $\partial C/\partial I$ may be large and an increase in Income tend to *raise* the investment rate of interest. In these circumstances, the situation is unstable at *any* given money rate; it is only an imperfectly elastic monetary system—a rising LL curve—that can prevent the situation getting out of hand altogether.

These, then, are a few of the things we can get out of our skeleton apparatus. But even if it may claim to be a slight extension of Mr. Keynes' similar skeleton, it remains a terribly rough and ready sort of affair. In particular, the concept of "Income" is worked monstrously hard; most of our curves are not really determinate unless something is said about the distribution of Income as well as its magnitude. Indeed, what they express is something like a relation between the price-system and the system of interest rates; and you cannot get that into a curve. Further, all sorts of questions about depreciation have been neglected; and all sorts of questions about the timing of the processes under consideration.

---

greater the elasticities of the CC and SS curves, and the larger is $\partial C/\partial I$ relatively to $\partial S/\partial I$. When $\partial C/\partial I > \partial S/\partial I$, the IS curve is upward sloping.

[9] Cf. Keynes, *General Theory*, p. 242.

[10] Cf. Myrdal, "Gleichgewichtsbegriff," in *Beiträge zur Geldtheorie*, ed. Hayek.

The *General Theory of Employment* is a useful book; but it is neither the beginning nor the end of Dynamic Economics.

J. R. Hicks

Gonville and Caius College
Cambridge

MONOPOLY OVER TIME[1]

By Gerhard Tintner

The greatest progress in economic theory in recent years has doubtless been made in the theory of monopoly. It is only necessary to mention the excellent work of Chamberlin and Mrs. Robinson on imperfect competition on the one hand, and the consideration of monopoly over time by Evans and Roos on the other.[2] What is missing seems to be a link between the classical and the modern approach to monopoly. It is the goal of this paper to develop some tools and concepts which may prove useful not only in the very simplified cases which are here considered, but also for questions like bilateral monopoly over time, imperfect competition over time, and so on.

It is the aim of this paper to consider at the same time the classical Cournot case of simple monopoly and the question of monopoly over time in various aspects. It is of course clear that these cases in themselves are rather exceptional and intrinsically of minor importance. Their significance, however, lies in the fact that their treatment enables one to develop concepts and notions which are very useful for the consideration of other more complicated and more significant types of monopoly. The difference between the simple Cournot case[3] and the case of monopoly over time lies in the demand curve. In the Cournot case the demand depends only on the price, in the other case it depends also on the rate of change of price over time and possibly higher derivatives.[4] The economic significance of this latter situation lies in the fact that the previous prices influence the buyers' reactions to the present price. This may have a very great importance, especially for goods that can be stored. The analysis of the problem of stocks, and of economic expectations and anticipations in general, is a very important one but it cannot be dealt with here. We simply assume in the Marshal-

[1] The author had the great pleasure of presenting the content of parts of this paper before discussion groups of the Economics Departments of Harvard University and the University of Chicago. He is much indebted to Professors J. A. Schumpeter and W. Leontieff at Harvard and H. Schultz and F. Knight at Chicago for valuable advice and criticism; also to Mr. M. Friedmann (Chicago), Mr. M. Millikan (Yale), and Mr. J. Ripley (Yale) for their kindness in correcting the English manuscript.

[2] E. C. Chamberlin, *The Theory of Monopolistic Competition*, Cambridge, Mass., 1933; J. Robinson, *The Economics of Imperfect Competition*, London, 1933; G. C. Evans, *Mathematical Introduction to Economics*, New York, 1930; C. F. Roos, *Dynamic Economics*, Bloomington, Ind., 1934.

[3] A. Cournot, *Researches into the Mathematical Principles of the Theory of Wealth*, New York, 1927, p. 56 ff.

[4] G. C. Evans, *op. cit.*, p. 143 ff.; C. F. Roos, *op. cit.*, p. 63 ff.

lian fashion[5] that the reactions of the buyers are pictured in the demand curve or surface.

There is, however, the possibility of linking up these ideas with utility theory by constructing a dynamic utility function. The utility is supposed to depend not only on the quantities possessed and, through the budget equation, on the existing prices (in equilibrium), but also on the expected prices. The expectation of prices will largely depend on the price tendencies. Thus the derivatives of prices with respect to time enter as parameters into the utility functions of the individuals and also into the demand functions herefrom derived. This reasoning links up the problems of monopoly over time with the very important and general questions of expectations, economic horizon, and the rôle of time in economic life in general.[6]

We are going to treat and compare the three following cases of monopoly: (a) the Cournot case, in which the demand depends only on the price at the given moment; (b) the Evans case, in which the demand depends on the existing price and the rate of change of the price; (c) the generalized case, in which the quantity demanded depends on the existing price, on the rate of change of price, and also on the higher derivatives of the price with respect to time.[7] This however is not the most general case conceivable, since demand functions involving integrals have been treated by Evans and Roos. The most general case would be a functional.[8] The results will be formulated: (1) in terms of profits, (2) in terms of revenue, and (3) in terms of demand. Our treatment requires reformulation of the classical theory of simple monopoly.

The following are the symbols we are going to use in our discussion: π denotes the profit, R the total revenue, C the total costs, p the price and D the quantity demanded, so that

$$\pi = R - C, \tag{1}$$

i.e., the profit is the difference between revenue and costs.

$$R = pD, \tag{2}$$

i.e., the revenue equals the price per unit times the quantity. Hence

$$\pi = pD - C. \tag{3}$$

[5] A. Marshall, *Principles of Economics*, 8th Ed., London, 1930, p. 92 ff.

[6] P. N. Rosenstein-Rodan, "The Rôle of Time in Economic Theory," *Economica*, new series, Vol. I, p. 77 ff.; O. Morgenstern, "Das Zeitmoment in der Wertlehre," *Zeitschrift fuer Nationaloekonomie*, Vol. V, p. 433 ff.; J. M. Keynes, *The General Theory of Employment, Interest and Money*, London, 1936.

[7] G. C. Evans, *op. cit.*, p. 154 ff.

[8] C. F. Roos, *op. cit.*, p. 14 ff.

In the Cournot case of simple monopoly the problem consists in making π a maximum. The monopolist tries to make his profit as big as possible by varying the price. This is expressed by

$$\frac{d\pi}{dp} = 0, \tag{4}$$

or, from (1),

$$\frac{dR}{dp} - \frac{dC}{dp} = 0. \tag{5}$$

Since

$$\frac{dC}{dp} = \frac{dC}{dD}\frac{dD}{dp}, \tag{6}$$

equation (5) can be written in the form

$$\frac{dR}{dp} - \frac{dC}{dD}\frac{dD}{dp} = 0, \tag{7}$$

or, from (2), as

$$D + p\frac{dD}{dp} - \frac{dC}{dD}\frac{dD}{dp} = 0. \tag{8}$$

We note that (4) expresses the condition for a maximum in the Cournot case in terms of profits, (7) in terms of revenue, and (8) in terms of demand.

We want now to treat the problem of monopoly over time in a similar manner. In the Evans case the quantity demanded depends not only on the existing price but also on the rate of change of price p':

$$D = D(p, p'). \tag{9}$$

This relation between demand, D, and rate of change of price, p', is by no means a simple one. For instance, a rising tendency of price may lead people to expect the price rise to continue or to expect that it may end soon. In the first case it would lead to an increased, in the latter to a decreased, demand. There is, however, reason to believe that the first case will be more frequent.[9]

By (1) the profit can also be considered as a function of p and p':

$$\pi = \pi(p, p'). \tag{10}$$

[9] G. C. Evans, *op. cit.*, p. 143. See also R. H. Whitman, "The Problem of Statistical Demand Techniques for Producers' Goods: An Application to Steel," *Journal of Political Economy*, Vol. 42, p. 577 ff.

The problem is to make the integral

$$\Pi = \int_{t_0}^{t_1} \pi(p, p')dt, \tag{11}$$

between two given points in time, t_0 and t_1, a maximum.

This is a problem of the calculus of variations. The necessary but not sufficient condition for a maximum is that the Euler equation is satisfied:

$$\frac{\partial \pi}{\partial p} - \frac{d}{dt}\frac{\partial \pi}{\partial p'} = 0. \tag{12}$$

In our case the variable t, or time, does not enter explicitly in the function π. That means that we assume the form of the demand surface constant over time. We therefore can give a first integral of the Euler equation in the following form:

$$\pi - p'\frac{\partial \pi}{\partial p'} = \pi_0, \tag{13}$$

where π_0, the constant of integration, is easily seen to be the value of π, or the profit, which would be made in the case of the simple Cournot monopoly. It is the integral of equation (4) as well as the result which would be obtained from (12) or (13) if the profit were independent of the rate of change of price, that is $\partial\pi/\partial p' = 0$. Equation (13) expresses the solution of the problem of monopoly over time in terms of profits.

In order to express it in terms of revenue we must again use equation (1). This gives us

$$\pi - p'\left(\frac{\partial R}{\partial p'} - \frac{dC}{dD}\frac{\partial D}{\partial p'}\right) = \pi_0. \tag{14}$$

To express our solution in terms of demand we use equation (2) and get

$$\pi - p'\left(p\frac{\partial D}{\partial p'} - \frac{dC}{dD}\frac{\partial D}{\partial p'}\right) = \pi_0. \tag{15}$$

This is the solution of monopoly over time in the simplest case. It is neccessary to add that, if the actual form of the demand surface is given, we can determine p and p' and, therefore all our variables, as functions of the time. The solution of (12) will then describe the most rational behaviour of the monopolist over time. Since (12) is a differential equation of the second degree, its solution contains two arbitrary constants. These can be determined by fixing the value of p for the

times t_0 and t_1 or by two other conditions. The problem is then entirely determinate.

The more general case of monopoly over time is analogous to the case just considered. The difference consists only in the fact that the quantity demanded, D, depends not only on the rate of change of price, p', but also on the higher derivatives of the price with respect to time:

$$D = D(p, p', p'', p''', \cdots). \tag{16}$$

The Euler equation has the form

$$\frac{\partial \pi}{\partial p} - \frac{d}{dt}\frac{\partial \pi}{\partial p'} + \frac{d^2}{dt^2}\frac{\partial \pi}{\partial p''} - \frac{d^3}{dt^3}\frac{\partial \pi}{\partial p'''} + \cdots = 0, \tag{17}$$

and its first integral, analogous to (13), is

$$\begin{aligned} \pi &+ p'\left(-\frac{\partial \pi}{\partial p'} + \frac{d}{dt}\frac{\partial \pi}{\partial p''} - \frac{d^2}{dt^2}\frac{\partial \pi}{\partial p'''} + \cdots\right) \\ &+ p''\left(-\frac{\partial \pi}{\partial p''} + \frac{d}{dt}\frac{\partial \pi}{\partial p'''} - \cdots\right) \\ &+ p'''\left(-\frac{\partial \pi}{\partial p'''} + \cdots\right) + \cdots = \pi_0. \end{aligned} \tag{18}$$

This solves the problem in terms of profits. The solution in terms of revenue is similar to (14):

$$\begin{aligned} \pi + p'\Bigg[&-\left(\frac{\partial R}{\partial p'} - \frac{dC}{dD}\frac{\partial D}{\partial p'}\right) + \frac{d}{dt}\left(\frac{\partial R}{\partial p''} - \frac{dC}{dD}\frac{\partial D}{\partial p''}\right) \\ &- \frac{d^2}{dt^2}\left(\frac{\partial R}{\partial p'''} - \frac{dC}{dD}\frac{\partial D}{\partial p'''}\right) + \cdots\Bigg] \\ + p''\Bigg[&-\left(\frac{\partial R}{\partial p''} - \frac{dC}{dD}\frac{\partial D}{\partial p''}\right) \\ &+ \frac{d}{dt}\left(\frac{\partial R}{\partial p'''} - \frac{dC}{dD}\frac{\partial D}{\partial p'''}\right) - \cdots\Bigg] \\ + p'''\Bigg[&-\left(\frac{\partial R}{\partial p'''} - \frac{dC}{dD}\frac{\partial D}{\partial p'''}\right) + \cdots\Bigg] + \cdots = \pi_0; \end{aligned} \tag{19}$$

and the solution in terms of demand, analogous to (15), is

$$
\begin{aligned}
&\pi + p'\left\{ -\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p'}\right] + \frac{d}{dt}\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p''}\right] \right.\\
&\qquad\qquad\qquad\qquad \left. - \frac{d^2}{dt^2}\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p'''}\right)\right] + \cdots \right\}\\
&+ p''\left\{ -\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p''}\right] \right.\\
&\qquad\qquad\qquad\qquad \left. + \frac{d}{dt}\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p'''}\right] - \cdots \right\}\\
&+ p'''\left\{ -\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p'''}\right] + \cdots \right\} + \cdots = \pi_0.
\end{aligned}
\tag{20}
$$

The solution will give the price and its derivatives as functions of the time. If the demand surface (16) depends on the first m derivatives as well as on the price itself, (17) is a differential equation of degree $m+1$ and we need $m+1$ conditions for the determinate solution.

In order to discuss and compare these results, we wish to transform them and to express them in terms of logarithmic derivatives. These have the advantage that they are pure numbers of dimension zero and therefore invariant with respect to a change in the scale. Since the scales, in our case the units of the quantities and of money, are arbitrary and have no economic significance, this seems to be of some advantage. In accordance with the usual procedure, the logarithmic derivative of any quantity with respect to price will be called elasticity.[10]

Let us now proceed to reformulate the results already reached for the case in which the solution is expressed in terms of profits. We get for the Cournot case, from (4),

$$\frac{d \log \pi}{d \log p} = 0. \tag{21}$$

This shows that the elasticity of profits must necessarily be zero if the profit is to be as large as possible.

In the Evans case we get, from (13),

$$\frac{\partial \log \pi}{\partial \log p'} = \frac{\pi - \pi_0}{\pi}. \tag{22}$$

This shows that the elasticity of profits with respect to the rate of change of price, p', varies inversely proportionally to the profit itself.

[10] A. Marshall, *op. cit.*, p. 102 ff. p. 838 ff.

Equation (22) represents a curve in the $\partial \log \pi/\partial \log p'$, π plane which indicates the relation between the elasticity of profits with respect to rate of change of price, $\partial \log \pi/\partial \log p'$, and the profit itself, π, which must be fulfilled at each point in time in order that the integral (11) should have a maximum. The hyperbola has two branches, the asymptotes are the lines $\partial \log\pi/\partial \log p'=1$ and $\pi=0$. If the profit is independent of the rate of change of the price, p', then the left-hand side of equation (22) is zero, and we get the simple Cournot case, the solution of equation (21).

In the general case, the solution is, of course, more complicated. We get, from (18),

$$\begin{aligned} & \frac{\partial \log \pi}{\partial \log p'} + \frac{\partial \log \pi}{\partial \log p''} + \frac{\partial \log \pi}{\partial \log p'''} + \cdots \\ & + \frac{p'}{\pi}\left[- \frac{d}{dt}\frac{\partial \pi}{\partial p''} + \frac{d^2}{dt^2}\frac{\partial \pi}{\partial p'''} - \cdots \right] \\ & + \frac{p''}{\pi}\left[- \frac{d}{dt}\frac{\partial \pi}{\partial p'''} + \cdots \right] + \cdots = \frac{\pi - \pi_0}{\pi}. \end{aligned} \tag{23}$$

The first series of terms is not difficult to interpret. It is the sum of the elasticities of profit with respect to p', p'', and so on. But the expressions in brackets are very difficult to explain. They are, however, first and higher order derivatives with respect to the time of the partial differential quotients of profit with respect to p'', p''', and so on. The right-hand side of the equation is the same as in (22). If the profit is independent of the time derivatives of the price, the left-hand side becomes zero and we get once more the solution of the simple Cournot case. The formal analogies are, however, greater between the Evans case and the Cournot case than between the latter and the general case.

We now proceed to give the solution in terms of revenues. In the simple Cournot case we get, from (7),

$$\frac{d \log R}{d \log p} = \frac{1}{p}\frac{dC}{dD}\frac{d \log D}{d \log p}. \tag{24}$$

This formula presents some interesting new formulations. The expression $d \log R/d \log p$ is nothing else than the elasticity of revenue with respect to price. This expression is not to be confused with the marginal revenue[11] or any magnitude derived from it. Marginal revenue means the change in the total revenue if the quantity is the independent variable rather than the price. The expression $d \log D/d \log p$ is of course the elasticity of demand. The remainder on the right-hand side

[11] J. Robinson, *op. cit.*, p. 51 ff.

equals $(1/p)$ (dC/dD) and means marginal costs divided by price; or marginal costs not for the unit of quantity but for the amount the unit of money will buy. The similarity to the weighted marginal utility (*ophelimité ponderée*)[12] is very striking. So the meaning of this formula is: The elasticity of revenue with respect to price must be equal to the weighted marginal costs multiplied by the elasticity of demand, or: At the monopoly point, the ratio of the elasticity of revenue to the elasticity of demand must be the same as the ratio of the marginal costs to the price.

In the Evans case we get, from (14),

$$\frac{\partial \log R}{\partial \log p'} = \frac{\pi - \pi_0}{R} + \frac{1}{p}\,\frac{dC}{dD}\,\frac{\partial \log D}{\partial \log p'}. \tag{25}$$

The discussion of this formula runs as follows: The left-hand side of expression (25) is the elasticity of revenue with respect to the rate of change of price, p'. The first expression on the right-hand side is inversely proportional to the total revenue, R. The second expression is the weighted marginal costs multiplied by the partial elasticity of demand with respect to the rate of change of price. The analogy with formula (24) is very striking, if we disregard the first term on the right-hand side. These are the conditions which must hold true at every moment of time for the magnitudes involved, in order that the integral (11) (the profit over a period of time, Π) should be a maximum.

In the general case the conditions are a little more complicated. We get, from (19), the solution in terms of revenue:

$$\begin{aligned}
&\frac{\partial \log R}{\partial \log p'} + \frac{\partial \log R}{\partial \log p''} + \frac{\partial \log R}{\partial \log p'''} + \cdots \\
&\qquad + \frac{p'}{R}\left[-\frac{d}{dt}\left(\frac{\partial R}{\partial p''} - \frac{dC}{dD}\,\frac{\partial D}{\partial p''}\right) \right.\\
&\qquad \left. + \frac{d^2}{dt^2}\left(\frac{\partial R}{\partial p'''} - \frac{dC}{dD}\,\frac{\partial D}{\partial p'''}\right) - \cdots \right] \\
&\qquad + \frac{p''}{R}\left[-\frac{d}{dt}\left(\frac{\partial R}{\partial p'''} - \frac{dC}{dD}\,\frac{\partial D}{\partial p'''}\right) + \cdots \right] + \cdots \\
&= \frac{\pi - \pi_0}{R} + \frac{1}{p}\,\frac{dC}{dD}\left(\frac{\partial \log D}{\partial \log p'} + \frac{\partial \log D}{\partial \log p''} + \frac{\partial \log D}{\partial \log p'''} + \cdots\right).
\end{aligned} \tag{26}$$

The interpretation of this somewhat more involved equation is not so easy. The first series of terms on the left-hand side, however, is the sum of the elasticities of the revenue with respect to the first and the

[12] V. Pareto, *Manuel d'Économie Politique*, Paris, 1909, p. 159 ff.

higher time derivatives of the price. The first term on the right-hand side is the same inversely proportional function of R as before. The rest are the weighted marginal costs multiplied by the sum of the partial elasticities with respect to the time derivatives of the price. The second series of terms on the left-hand side are derivatives with respect to the time of the partial derivatives of revenue with respect to p'', p''', and so on.

The next step is the formulation of our results in terms of demand. From equation (8), this leads, in the Cournot case, to the well-known formula:[13]

$$\frac{d \log D}{d \log p} = \frac{p}{\dfrac{dC}{dD} - p}. \tag{27}$$

The elasticity of demand must equal the ratio of the price to the difference between marginal costs and price. This result can be reformulated in a slightly different manner:

$$\frac{d \log D}{d \log p} = \frac{R}{\dfrac{dC}{dD} D - R}. \tag{28}$$

This formula means: At the monopoly point, the elasticity of demand equals the ratio between the revenue and the difference between two magnitudes. The first is the quantity demanded multiplied by the marginal costs. It represents the amount of money which would be paid by the consumers if, under free competition, the equilibrium were established at the same point. Under free competition the price equals the marginal costs. This expression does not, in contrast to the concept of consumers' rent, involve a comparison of utilities of different individuals.[14] The hypothesis is that the quantity which is produced under monopoly should be the same as that produced under free competition. The second quantity in the formula is again the revenue realised by the monopolist.

The formulation in the Evans case is very similar. We get, from (15):

$$\frac{\partial \log D}{\partial \log p'} = \frac{\pi - \pi_0}{D\left(p - \dfrac{dC}{dD}\right)}; \tag{29}$$

[13] L. Amoroso, "La Curva Statica di Offerta," *Giornale degli Economisti*, 1930, p. 11 ff. See also A. P. Lerner, "The Concept of Monopoly and the Measurement of Monopoly Power," *Review of Economic Studies*, Vol. I, p. 161 ff.

[14] Cf. L. Robbins, *An Essay on the Nature and Significance of Economic Science*, 2nd Ed., London, 1935, p. 139 ff.

and this means: The partial elasticity of demand with respect to the rate of change of price must equal the difference between the profit actually realised and the profit which the monopoly would yield in the simple Cournot case, divided by the difference between price and marginal costs multiplied by the quantity actually sold. Analogously to (28) this can be expressed as follows:

$$\frac{\partial \log D}{\partial \log p'} = \frac{\pi - \pi_0}{R - \dfrac{dC}{dD} D}. \tag{30}$$

The partial elasticity of demand with respect to the rate of change of the price must equal the following quantity: The difference between the profit actually realised and the profit in the Cournot case, divided by the difference between the revenue actually realised and the amount of money which the consumers would pay, if this were an equilibrium point under conditions of free competition. We get an expression in terms of marginal costs:

$$\frac{\partial \log D}{\partial \log p'} = \frac{\pi - \pi_0}{R\left(1 - \dfrac{1}{p}\dfrac{dC}{dD}\right)}. \tag{31}$$

The partial elasticity of demand with respect to the rate of change of the price equals the difference between the profit actually made and the profit which would be made in the Cournot case, divided by the revenue multiplied by one minus the weighted marginal costs.

If we try to express these results for the general case the situation is somewhat more complicated. We get, from (20):

$$\begin{aligned}
&\frac{\partial \log D}{\partial \log p'} + \frac{\partial \log D}{\partial \log p''} + \frac{\partial \log D}{\partial \log p'''} + \cdots \\
&\quad + \frac{}{D\left(p - \dfrac{dC}{dD}\right)}\left\{- \frac{d}{dt}\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p''}\right]\right. \\
&\quad \left. + \frac{d^2}{dt^2}\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p'''}\right] - \cdots\right\} \\
&\quad + \frac{p''}{D\left(p - \dfrac{dC}{dD}\right)}\left\{- \frac{d}{dt}\left[\left(p - \frac{dC}{dD}\right)\frac{\partial D}{\partial p'''}\right] + \cdots\right\} \\
&\quad + \cdots = \frac{\pi - \pi_0}{D\left(p - \dfrac{dC}{dD}\right)}.
\end{aligned} \tag{32}$$

The first series of terms on the left-hand side represents the sum of the partial elasticities of demand with respect to the first and the higher time derivatives of the price. On the right-hand side we have the difference between the profit actually realised and the profit in the Cournot case. This quantity is divided by the difference between the price and the marginal costs multiplied by the amount sold. The remaining terms are again time derivatives of the partial differential quotients of demand with respect to p'', p''', and so on.

In the present paper we have endeavored to consider three different cases of monopoly: (a) the Cournot case, (b) the Evans case, (c) the more generalised case. It is easy to see from the formulae developed that both the Cournot case and the Evans case are special solutions of the general case.

We have examined the solutions or the conditions of maximum monopoly profit from three different points of view: (1) profits, (2) revenue, (3) demand. Our formulation is in terms of what we would be tempted to call economic parameters, i.e., magnitudes that have either a definite economic meaning attached to them from previous economic discussions or are formed analogously to them, elasticities of demand with respect to the price and the rate of change of the price, marginal costs, and weighted marginal costs. It was pointed out that in the Cournot case and in the Evans case the formulation was possible in those terms. In the general case however the solution was not so simple, and it is doubtful if much economic meaning can be attached to the results reached. The analogy between the formulations in the Cournot and in the Evans case is striking and gives some hope that, at least in the less complicated cases of monopoly over time, the treatment in economically significant terms may not be impossible.

It should be mentioned that we have left second order conditions out of consideration. The question whether a maximum or a minimum will be realised is unimportant in the case of monopoly since, at worst, the monopolist would rather produce nothing at all than incur a loss. It is interesting, however, if treated from the point of view of utility theory.[15]

Gerhard Tintner

Cowles Commission for
Research in Economics
Colorado Springs, Colorado

[15] H. Hotelling, "Demand Functions with Limited Budgets," Econometrica, Vol. 3, Jan., 1935, p. 71 ff.

THE SENSITIVITY OF TAXES TO FLUCTUATIONS* OF TRADE

By R. F. Bretherton

I. GENERAL THEORY

The subject of the sensitivity of the yield of various types of taxes to the movements of the trade cycle is of great practical importance; but it has been almost completely overlooked in the standard English works on public finance[1]—perhaps because, until recently, the science of public finance has been built almost entirely upon a framework of theories of a static or evenly progressive economic system. It therefore seems worth while to outline a theory of the subject, and to try to obtain some simple statistical measurements.

The first problem is to discover a satisfactory standard against which to measure the "cycle sensitivity" of taxes, which can be applied at different times and in different countries. Since money income is the ultimate source from which all taxes must be paid, it seems desirable to measure fluctuations in the yield of particular taxes against fluctuations in the total of national social income,[2] wherever satisfactory figures for these are available. It is, of course, previously necessary to eliminate the effects upon the yield of alterations in the rates or basis of the tax, which have taken place during the period under review. We can then say that the cycle sensitivity of any tax, in any year, or in any phase of the trade cycle, is measured by the formula

$$\frac{\text{Percentage Change in Yield}}{\text{Percentage Change in National Social Income}}\,.$$

This method eliminates from the result the effects of general price movements, and of the expansion or contraction of population; though of course it leaves untouched the effects of long-period changes of relative demand and of the distribution of wealth, in so far as these are independent of the trade cycle. Satisfactory elimination

* Paper read at the meeting of the Econometric Society, Oxford, England, September, 1936.

[1] There is no explicit discussion of the subject in any of the standard English works; and in America, even so recent a book as C. L. King's *Public Finance*, 1935, makes no mention of it. There is much illustration of the effects of tax sensitivity in Dalton, *Unbalanced Budgets*, and in B. Thomas, *Monetary Policy and Crises—A Study of the Swedish Experience*. See also a brief discussion of the subject by H. Laufenburger in "The Interaction of Taxation and Trade Fluctuations," *Index*, May, 1935.

[2] National Social Income may be briefly defined as the "aggregate of individual and collective incomes, less incomes received by compulsory reductions from other incomes in return for no services or services not rendered in the year in question." Bowley and Stamp, *The National Income, 1924*, p. 41.

of these trends, at least in recent years, does not appear to be possible on the basis of the existing statistical material; nor is it certain that it is desirable, since the character of the long-period trends is likely to be closely connected with the character of the trade cycle. In the theoretical discussion which follows, such changes are assumed to be nonexistent, except where specific mention is made of them.

As a starting point, we may consider the question of income taxes. A universal and strictly proportional income tax would clearly show unitary sensitivity, except in so far as the onset of a trade depression caused delays in collection or made it necessary to write off a larger proportion of assessments as uncollectable. There would, of course, be a lag in the timing of changes in tax yield and changes in the national income if the tax collected in one year was the result of assessments upon incomes earned in the previous year. A degressive income tax, with a low exemption limit and assessment below the standard rate on a range of income above it, would show a sensitivity slightly greater than 1; for unemployment, wage and salary cuts, and a loss of dividends to small investors would bring many incomes below the exemption limit in times of depression, which would come above it in prosperity. A fully progressive income tax would show a large positive sensitivity. This would be true even if personal incomes of every size were reduced proportionately in depression and increased proportionately in prosperity; for a general scaling down must reduce the amount of income assessable at high rates. It is, however, probable that in general the incomes of the very rich are in fact reduced by depressions and expanded by booms in a *greater* proportion than those of the moderately well to do[3]—though not, perhaps, always more than the average incomes of wage earners, if the effects of unemployment are allowed for. This, of course, very much increases the sensitivity of a progressive income tax; for it makes it a function both of its scale of graduation and of the degree of relative redistribution of income in the varying phases of the trade cycle.

Death duties are also likely to show a high degree of sensitivity. The numbers of estates which become liable to duty is not, indeed, a function of the state of trade (except in so far as trade depressions drive millionaires, like Kreuger, to suicide); but their valuation is very much affected by the level of stock exchange prices, which certainly fluctuates much more than the national income. And, in so far as the duties are progressive, changes in their yield will be more than proportionate

[3] In Great Britain, the total of incomes between £2,000 and £6,000 fell by 19.2% between 1929–30 and 1932–3; those between £6,000 and £50,000 by 35.1%; and those above £50,000 by 48.5%. (Computed from super-tax statistics given in the *78th Report of the Board of Inland Revenue.*)

to the changes in the total value of estates liable. But there is also another matter to be considered. The owners of large and very large fortunes are likely to have a greater proportion of their capital assets in the form of industrial equities, whose prices are more sensitive than those of gilt-edged securities or house property;[4] and this clearly increases the sensitivity of the yield. On the other hand, the prices of industrials and of gilt-edged move at different times and often in different directions. A rise in the prices of gilt-edged may take place while the trend of national income is still downwards, and industrials are still falling; and this would clearly reduce the sensitivity of death duties as a whole to the depression. There must also be considered the effects of long-period changes in the rate of interest and in general mortality; in the years since the War, these have tended to raise the importance of death duties as a source of revenue.

The miscellaneous group of taxes upon transfer of property known collectively as the Stamp Duties must always be highly sensitive; for both the number of transfers of property and (where this is relevant to the amount of the tax, which is not always) the value of the property transferred, are likely to be greater in prosperity than in depression. In particular, an active Stock Exchange must show itself in an increase of the revenue from transfers of shares, as well as from the duty on the share capital of new issues. The timing of the fluctuations of such duties on the transfer of property is, however, likely to be in advance of changes in the national income.

Taxes upon commodities must vary in sensitivity according to the type of commodity taxed, and according to the technical form of the tax. It is obvious that specific duties will show less sensitivity than those which are levied *ad valorem*, since they will not reflect directly changes in the price of the commodity. Against this, however, must be set the consideration that an unchanged *rate* of specific duty means a smaller percentage change in price to consumers if world prices alter, than does an unchanged rate of *ad valorem* duty. In the case of an article for which consumers' demand is very elastic, and remains elastic even though the position of the curve alters, this may be important in reducing the superiority of the specific duty as a method of maintaining a stable revenue. The types of commodity taxed may be divided in-

[4] In Great Britain, the total capital value of estates becoming liable to Estate Duty rose by 24.6% from 1922–3 to 1929–30, fell by 13.2% to 1931–2, and rose again by 14.6% to 1934–5. But the changes in the totals for great estates were larger than for small. For estates of less than £100,000, the changes were +31.0%, −8.4%, +10.6%; for estates of more than £100,000, +10.7%, −26.2%, +26.2%. Accidents of mortality among the very rich may have influenced these figures somewhat; but the greater sensitivity of large estates in the recent depression and recovery seems to be clear.

to two broad classes—viz., finished commodities, or raw materials entering into their manufacture, which are purchased by final consumers; and investment goods, or raw materials which are primarily important for their production. Since final consumption almost always varies less than the whole national income, we may expect taxes on consumption goods as a class to show a low sensitivity; and just the reverse will be true of investment goods. But it is important to make a subordinate distinction between goods for immediate consumption, and "consumption capital"; for experience shows[5] that, although the sales of even non-essential "immediates" such as beer and tobacco do not change rapidly in the course of the cycle, sales of such things as furniture and clothes may show a sensitivity almost as high as that of machinery and other investment goods proper. Further, it is necessary to stress the point that taxes on goods, for which the secular trend of demand is steeply upwards, because of the spread of new habits, may show little sensitivity to depression though a good deal to recovery; people acquire motor cars mainly in prosperity, but do not readily surrender them in adversity. On the other side, a period of depression may cause a rapid fall, with no subsequent recovery, in the yield of a tax on an article for which the secular trend of demand is downwards. Import and export duties in general may be expected to show greater sensitivity than excise duties on articles manufactured at home, because of the severity of trade-cycle influences upon international trade.[6] Finally, it must be noticed that where import duties are protective or preferential, their yield may be much affected by changes in the relation of national price structures and exchange rates, which are common, though in detail unpredictable, incidents of the trade cycle. An isolated depreciation of currency by taxing country will, by increasing the amount of protection afforded to home producers, usually diminish the yield of protective import duties; and if the depreciation is followed by some other countries, but not by all (as was the case with the depreciation of sterling in 1931), the resultant redistribution of the sources of imports may affect the yield of duties of a preferential character.

II. THE SENSITIVITY OF BRITISH TAXES SINCE 1922

Actual measurement of the sensitivity of taxes, even in one cycle, is not very easy. In the first place, no attempt can be made at present to eliminate the effects of long-period changes in relative demand and in the distribution of wealth, which are independent of the trade cycle. Second, no figures for the annual fluctuations of the national income are available for years before 1924; and those after that date are sub-

[5] Cf. Clark, "Strategic Factors in the Business Cycle."

[6] See League of Nations, *World Economic Survey, 1933–4*, pp. 230–1.

ject to an unknown, though probably small, margin of error.[7] Third, there are serious difficulties involved in computing an index of tax yield, where the rate or basis of the tax has been altered several times during the period; and of the more important British taxes, only one, the duty on spirits, has remained substantially unchanged both in rate and basis from 1922 to 1936. In the rough calculations which follow, the rough and ready expedient has been usually adopted of deducting from or adding to, the "actual" yield, the Budget estimate and the subsequent Treasury computation of the change of revenue caused by such alterations, in order to obtain the "hypothetical yield" of the tax at an unchanged rate throughout. In the case of income tax, sur-tax, and estate duties, a rough check in this method can be obtained, as far as changes of rate only are concerned, with the aid of the published statistics of assessable income and of the value of estates liable to death-duty.[8] On this latter basis, it is of course possible to calculate exactly the final products of the assessments made to those taxes; but because of the delays in assessment and collection, these do not correspond with the net annual receipts of revenue. It is recognized that where alterations in rates or bases of taxes have been numerous, or took place early in the period, the calculated "hypothetical yield" is subject to a considerable margin of error; but the results should be accurate enough to afford a general picture of the variations in behaviour of the receipts from the major British taxes during the period.

Indices of the "hypothetical yields" of the ten most important taxes are given in Table I for the years from 1921–2 to 1935–6, a period which covers the whole of one trade cycle, measured from one point of minimum activity to another, and three years of recent recovery in addition. In Table II the magnitude of the fluctuations of National Social Income and of the hypothetical yield of the various taxes is summarized for each phase of the cycle, an index of "cycle sensitivity" is computed, and the time relation of changes in tax yields to changes in national social income is indicated. In the computation of sensitivity, the total fluctuation in tax yield is divided by the total fluctuation in national social income in each phase; but periods do not necessarily correspond either in length or timing: for example, the yield of the Entertainments Tax rose from a "low" of 93.1 in 1923–4 to a "high" of 123.1 in 1930–1, while the National Social Income rose from its lowest of 95.5 in 1922–3 to its highest of 108.6 in 1929–30; but the sensitivity is computed on the basis of the total movement in each case. (30.0/13.1 =2.29.)

[7] The indices used are based, as regards the years from 1924–5, upon material, as yet unpublished, very kindly supplied by Mr. C. G. Clark; as regards the years 1922–3 and 1923–4, upon estimates worked out by the writer.

[8] Reports of the Board of Inland Revenue.

TABLE I

INDICES OF HYPOTHETICAL YIELDS OF BRITISH TAXES, 1921 TO 1935

1924–5 = 100

Financial Year	Income Tax	Super-tax and Sur-tax	Estate Duty: Average Yield per estate liable	Stamp Duties	Enter-tainments Tax	Motor Vehicle Duties	Tobacco: Customs and Excise	Sugar and Molasses: Customs and Excise	Beer: Customs and Excise	Spirits: Customs and Excise	National Social Income (Calendar Years)
1921–2	.....	96.0	91.2	85.6	.....	67.2	106.3	88.6	.....	.....	100.8
1922–3	91.9	102.3	103.4	96.0	98.9	76.4	102.3	105.9	92.3	105.1	95.5
1923–4	92.2	97.7	104.8	94.8	93.1	88.8	99.9	96.4	95.2	105.8	96.5
1924–5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1925–6	103.3	127.1	98.7	110.3	101.1	112.0	103.0	103.0	100.5	97.8	108.4
1926–7	96.5	132.4	95.4	109.3	101.4	121.4	103.8	100.8	102.6	85.2	102.4
1927–8	100.1	118.6	104.6	119.0	108.3	136.3	106.1	100.3	101.5	92.8	107.1
1928–9	100.3	111.1	108.4	133.6	106.3	138.0	105.2	94.7	92.4	89.5	107.4
1929–30	101.6	112.2	94.5	112.3	118.5	146.0	112.3	93.3	93.9	83.4	108.6
1930–1	100.2	118.9	99.3	91.8	123.1	154.0	114.8	100.3	89.0	79.6	106.0
1931–2	98.6	120.2	66.3	76.8	121.6	151.0	113.0	108.3	82.6	68.4	94.0
1932–3	76.2	86.8	77.1	85.1	120.7	155.2	113.1	94.4	74.0	67.6	92.3
1933–4	74.6	75.4	90.5	103.4	119.2	166.9	113.7	92.3	73.0	65.5	95.3
1934–5	78.5	74.5	85.1	114.4	127.8	175.4	119.7	95.2	75.6	63.4	101.5
1935–6	85.2	74.9	.....	121.3	129.8	188.4	128.3	92.3	78.2	68.5	107.7

(1) *The Income tax* shows the surprisingly low sensitivity of 0.74 to the upswing of the cycle. This may be partly due to a statistical bias introduced by large payments of arrears in the early years, and by complications arising from the system of three-year averaging for assessment under Schedule D, which was in force until 1927. But it is

TABLE II
SUMMARY OF TAX SENSITIVITIES, 1921–1935

	Movement of Index of Hypothetical Yield			Sensitivity Index*			Timing†		
	Upswing: 1st Lowest to Highest	Downswing: Highest to 2nd Lowest	Recovery: 2nd Lowest to 1936	Upswing	Downswing	Recovery	1st Lowest	Highest	2nd Lowest
National Social Income	+13.1	−16.3	+15.4	1.00	1.00	1.00	1922	1929	1932
Income Tax	+ 9.7‡	−27.0	+10.6	0.74	1.66	0.69	equal	equal	1 late
Super-Tax	+23.5§	−45.7	+ 0.5	1.79	2.80	0.003	1 late	2 late	2 late
Estate Duty, 1894	+17.2	−42.1	......	1.31	2.60	.....	1 early	1 early	1 early
Stamp Duties	+48.0	−56.8	+44.8	3.66	3.48	2.91	1 early	1 early	1 early
Entertainments Tax	+30.0	− 3.9	+10.6	2.29	0.24	0.69	1 late	1 late	1 late
Motor Vehicle Duties	+86.8	− 3.0	+37.4	6.63	0.18	2.43	1 early	1 late	1 early
Tobacco Duties	+14.9	− 1.8	+15.3	1.14	0.11	1.00	1 late	1 late	1 early
Sugar Duties	+19.7	−16.0	......	1.49	0.98	.....	1 early	2 late	2 late
Beer Duties	+ 1.6‡	−20.9	+ 5.2	0.12	1.28	0.34	equal	equal	1 late
Spirit Duties	−15.6‖	−26.1	+ 5.1	−1.19	1.60	0.33	equal	1 early	2 late

* Sensitivity Index = Movement of Index of Tax Yield, divided by movement of Index of National Social Income, in each phase of the cycle.

† Timing indicates the time relation in years of "turning points" of tax-yield to "turning points" of National Social Income.

‡ 1929–30 taken to be highest.

§ 1923–4 taken as 1st lowest, and 1931–2 as highest.

‖ 1928–9 taken to be highest.

also probable that the yield of income tax was affected by a secular tendency towards a redistribution of income in favour of the less wealthy classes of the community, which was the result of social policy rather than the trade cycle. For the depression, however, the sensi-

tivity was 1.66, and for recovery 0.69; the full force of the depression showed itself a year late, and the timing of recovery was similarly delayed.

(2) *Super-tax and Sur-tax* shows curious vagaries owing to extraordinary administrative delays in assessment and collection, the nature of which is brought out by the figures in Table III.[9] These delays caused the payment of the tax on income received in the boom years of 1919 and 1920 to be paid as much as four or five years later; and the

TABLE III
SUPER-TAX AND SUR-TAX, 1920–1934

YEAR	A Actual Net Receipts	B Net Tax Finally Assessed, and Due for Payment in Corresponding Year	C Excess (+) or Deficiency (−) of A over B	D Index of Hypothetical Net Tax Assessed	E Index of Hypothetical Net Receipts
	Millions of £	Millions of £	Millions of £	(1924–5 =100)	(1924–5 =100)
1920–1	55.669*	71.085	−15.416	104.2	.....
1921–2	61.351*	75.460	−14.104	111.1	96.0
1922–3	63.910	63.947	− 0.037	93.9	102.3
1923–4	61.747	66.833	− 5.086	98.2	97.7
1924–5	62.989	67.781	− 4.792	100.0	100.0
1925–6	67.833	57.681	+10.152	100.2	127.1
1926–7	66.295	57.064	+ 9.231	99.1	132.4
1927–8	60.053	55.073	+ 4.980	95.7	118.6
1928–9	56.214	56.125	+ 0.089	97.5	111.1
1929–30	56.624	59.691	− 3.367	103.7	112.2
1930–1	67.657	70.973	− 3.316	101.6	118.9
1931–2	77.083	68.342	+ 8.741	88.1	120.2
1932–3	60.310	55.186	+ 5.126	69.4	86.8
1933–4	52.395	48.350	+ 4.040	60.0	75.4
1934–5	50.916	45.770	+ 5.146	56.8	74.5

* Including Southern Ireland.

This table is designed to show the effects of administrative delays in assessment and collection upon the yields of Super-tax and Sur-tax. For example, Super-tax assessed for the year 1920–21 (on the incomes of 1919–20) ultimately amounted to 71.085 millions; but the actual receipts for that year amounted only to 55.669 millions. From 1925–6 to 1928–9, however, the actual receipts largely exceeded the amounts finally assessed, because of the payment of arrears.

[9] Until 1928–9, "Super-tax" was assessed on the income of the previous year, and fell due for payment on January 1st of the (financial) year of assessment. By the Finance Act, 1927, "Sur-tax" was instead to be assessed on the income of the year of assessment, and paid on January 1st. of the succeeding year, the payment being termed a "delayed installment of Income-tax." In practice, both assessment and payment are commonly much more "delayed" than the law allows.

same cause explains the fact that the later "peak" of the yield was delayed until the year 1931–2, when the depression was almost fully developed. A glance at the index of the amounts of Super-tax and Surtax finally assessed (Table III), which is free from these complications, shows that the tax had a lower sensitivity to prosperity even than the Income tax; this is again probably attributable to the effects of a redistribution of income independent of the trade cycle.[10] For the downswing of the cycle, the sensitivity was very high—2.80 in the case of annual receipts, and 2.88 in the case of tax ultimately assessed; and recovery came very late and has so far been slight.

(3) *Estate Duty* at first sight appears to behave in a random fashion; but when the effects of changes in general mortality have been eliminated by taking the hypothetical yield per estate (instead of all estates) as an index, it shows a distinct cyclical sensitivity, of 1.31 for the rise and 2.60 for the down-swing.[11] The timing was one year in advance in each case. As in the case of Income and Super-tax, the sensitivity was probably distorted by changes in the distribution of wealth.

Thus all the three main direct taxes show a high cycle sensitivity. In 1924–5 they provided between them 56% of the tax revenue; by 1934–5, in spite of numerous increases in their rates, the proportion had fallen to 49%.

(4) *Stamp Duties* show a most perfect cyclical pattern, though as might be expected, they move a year in advance of the national income. With the exception of slight checks in 1923–4 and 1926–7, there was a steady rise to 1928–9 (sensitivity 3.66); an equally sharp fall to 1931–2 (3.48); and a large recovery to 1935–6. (2.91).

(5) *Entertainments Tax* appears to show considerable sensitivity to improving conditions, but very little to depression. It is levied upon a service for which the trend of demand is certainly upwards; and the cinema habit appears to be one which, once acquired, is not easily surrendered. There was a practically steady rise to 1930-31 (sensitivity 2.29); a slight fall for three years (0.24); and a sharp recovery in 1934–5 and 1935–6 (0.69). There was a lag of one year in the timing throughout.

[10] Of incomes accruing in 1924–5, and liable to Super-tax, 44.3% belongs to persons with incomes between £2,000 and £6,000, 48.7% to persons with incomes between £6,000 and £75,000, and 7% to persons with incomes of over £75,000. In 1929–30 the percentages were 47.0, 46.4, and 6.6. Incomes below the level of £25,000 increased in number, and in total from £452 million to £491 millions; but the number of incomes above £25,000 decreased, and their total amounts fell from £109 millions to £102 millions. With progressive rates of assessment, this of course very much damped down the increase of yield.

[11] Though it is possible by this device to eliminate the effects of *general* mortality, the figures are still somewhat distorted by accidental coincidences of deaths among the very rich.

(6) *Motor Vehicle Duties*, even more than the Entertainments tax, were borne up on a wave of expanding demand, which was almost sufficient to obliterate the influence of the trade cycle. Sensitivity to the rise was 6.63; to the fall 0.18; and to recovery 2.43. There was a lag of one year at the peak, and an advance of one year at the depression. But though the movement was thus almost continuously upward, changes in the national income are fairly closely related with changes in the rate of growth of the yield of the duties, except in the one year 1930–1.

(7) *Tobacco Duties* show a gently rising trend, with very low cycle sensitivity. The peak was reached in 1930–1 (sensitivity 1.14); and for the next four years the curve was almost level; there was then a moderate recovery in 1934–5, and a more considerable one in 1935–6. (Sensitivity, 1.00.) This is probably a very typical pattern for a final consumers' good of this character.

(8) *Sugar and Molasses Duties*. Here the repeated increases of the customs preference to Imperial sugar, combined with the expansion of the production of home-produced beet sugar, make it difficult to calculate the hypothetical yields with any reasonable accuracy; the fluctuations in the curve are mainly due to variations in the proportions coming from Imperial and foreign sources. The steady increase in the proportion of Imperial sugar was checked in 1930 and 1931 by an alteration in the prices of sugar from foreign and Imperial sources; this was corrected later by the depreciation of the pound.[12] The curve of "actual" yield, with its fall from an index of 93.5 in 1925–6 to 57.3 in 1934–5, gives some idea of the loss of revenue which has resulted from the policy of preference and of the instability which it has introduced into the yield of this tax; for the rates of duty on foreign sugar have remained unchanged since the Budget of 1924.

(9) *Beer Duties* were certainly subject to a strong downward trend. On the whole, they show a slight positive sensitivity to 1929-30, a sensitivity of 1.28 to the depression, and one of 0.34 to recovery.

(10) *Spirit Duties*, the rates of which remained unchanged throughout the period, show an even steeper downward trend, with apparently a low sensitivity to the trade cycle, apart from an acceleration of the rate of fall in 1931–2. They had shown no sensitivity to the recovery by 1934–5; but there was a sharp rise in the following year.

For most purposes, however, the sensitivity of particular taxes is less important than the average sensitivity of the whole system; the art of

[12] See *Economist, Budget Supplement*, 1935. Also the Reports of H. M. Commissioner for Customs and Excise. The ratio of Empire to total retained imports of sugar was 22.5 in 1925–6, 38.0 in 1928–9, and 46.6 in 1929–30. It fell to 33.3 and 33.4 for the next two years, then rose to 43.6 and 51.6 for 1932–3 and 1933–4.

successful budgeting consists in allowing extremes to cancel out. The ten taxes so far examined provided 93.5% of the total tax revenue of the British government in 1924–5; so they may be regarded as representative of the whole system as it was at that date. When taken together, their yields showed a sensitivity of 0.69 to the upswing, from 1921–2 to 1927–8; a considerably higher figure of 1.18 for the downswing from that year to 1933–4 (though the decline was not really serious before 1931–2); and a low sensitivity of 0.53 to recovery to 1935–6. On the whole, this appears to be only a moderate degree of sensitivity; but it is worth considering how it was attained. Considerable relief came from the unevenness of the "timing" of fluctuations of particular taxes; Stamp Duties and Estate Duty began to fall long before Income tax and Super-tax had reached their peak, and had made an appreciable recovery before the taxes on income touched bottom. Had the years 1927 to 1929 been years of real trade boom, instead of merely moderate prosperity, this smoothing effect would probably have been much more noticeable. There was also a fortunate balancing of taxes such as Entertainments tax and Motor Vehicles Duties, in which the secular trend was steeply upwards and sensitivity to depression small, against others, like the Beer and Spirit Duties, in which the long-period trend was downwards, and sensitivity to depression considerable. The writer has not examined in detail the sensitivities of continental or American tax systems; but superficial inspection suggests that they showed much greater sensitivities than that of Great Britain; and this must certainly be partly responsible for the greater severity of budget difficulties in many of those counties since 1929. But even in Britain, the sensitivity of the tax system was large enough to cause serious trouble, because it was considerably greater than the sensitivity of expenditure. On the basis of unchanged rates, the hypothetical yield of these ten taxes fell from 90.6% of total tax revenue in 1927–8 to 72.3% in 1933–4; while expenditure by the central Government (other than that for the redemption of debt)[13] rose slowly but fairly steadily from a first low point of £748 millions in 1923–4 to £781 millions in 1929–30; it then jumped sharply to £833 millions in 1932–3, fell to £770 millions in the following year and then began to rise again. In so far as it was not met by borrowing, the gap had to be filled, either by increases of rates, of by the invention of fresh taxes. Broadly speaking, about one half of the gap was met by increasing the rates of old taxes and the other half by developing new fields of taxation, mainly in the tariff.

[13] Exchequer Issues *less* New Sinking Fund. These figures exclude the deficit on the Unemployment Insurance Fund before 1932–3, and also all Capital Expenditure not chargeable against revenue.

III. SOME CONSIDERATIONS OF POLICY

The theory and facts so far set out seem to make it necessary to add another canon or criterion of a tax system to the familiar list—namely, its cyclical sensitivity. But it is not altogether obvious that it is an evil to have a tax system which is highly sensitive to the trade cycle.The adherents of "orthodox principles" of public finance would indeed argue that this must lead either to unbalanced budgets or to rapid alterations in the rates or basis of taxation; since both these are regarded as evils, the ideal situation is one in which the yield of existing taxes responds automatically to fluctuations of public expenditure. Now, since, on the whole, in a modern state with well-developed social services, public expenditure tends automatically to expand in bad times, if not to contract in good (quite apart from changes in the policy of expenditure, which may be induced by the cycle), the principle of stability of rates would dictate the selection of taxes which are *inversely* sensitive to the Trade Cycle—if, indeed, they can be found. Failing that, we should at least concentrate on taxes which show only a small positive sensitivity. In practice, this would mean large use of taxes either on articles of general consumption, or on goods and services for which the secular trend of demand is upwards. But the first expedient would conflict with almost any principle of equity, since it would increase the burdens on the poor at times when they were least able to bear them; and the second is of doubtful advantage from the point of view of trade-cycle policy. When there is a general falling off of investment, it is desirable to stimulate the development of new demands, rather than to retard them by heavy taxation. Thus the refusal of the British government to increase the Motor Vehicle Duties along with other taxes in 1931 was probably good policy, as was their actual reduction of them at an early stage in the recovery in 1934. But the hole in the Budget caused by the cycle sensitivity of other taxes was made good largely by the imposition of a tariff; this may have stimulated recovery in some directions, but it certainly increased both the relative and absolute burdens on the poor. There appears to be a necessary conflict between the canon of equity and the canon of stability. Some relief might perhaps be given by retaining the emphasis on the sensitive direct taxes, but by assessing them on the average of several years' income, as was at one time partly done with the Income tax, and by "staggering" the dates of collection, as happens to some extent already with Income tax and Sur-tax. In any case, careful attention should be paid to the timing of the fluctuations of various taxes.

The adherents of less orthodox methods of budgeting in the trade cycle might at first thought believe that a highly sensitive tax system is not a bad thing, if only no attempt is made to meet the resultant def-

icits by additional taxation or to dispose of surpluses by reducing the rates. Let us meet our deficits by borrowing in depression, and use the surpluses of the boom to pay off the resultant debts; and there will be no need to alter the form or the rates of existing taxation at all. The more sensitive the tax system, the more easily can the State expand credit in depression and contract it in boom; and the more, too, will the amount of its revenue depend on the success of its own trade-cycle policy. But even if we accept this general view of the functions of the budget as an instrument of trade control, this particular argument seems to be too simple. To combat a depression which has already developed, it may be desirable actually to *reduce* the rates of some forms of taxation; for the psychological effect of a reduction of rates is much greater than the mere knowledge that they will not be raised. Further, as was argued above, there may be a special case for giving the stimulus of a reduction of taxes to commodities and services of expanding demand, the yield of which is comparatively insensitive to the depression. On the other hand, it may even be desirable to increase the rates of such direct taxes as fall mainly upon the rich; for in depression the State will be a more active spender and investor than wealthy private persons. Similarly, variations of rates of taxation may be a useful instrument for the control of the upswing of the cycle; increased taxation of stock exchange transactions and of the products of new industries may steady the rate and balance the direction of new investment, while reductions of direct taxation may stimulate saving—supposing, of course, that this is admitted to be *sometimes* conducive to prosperity! Control of this kind in the upswing is rendered all the more necessary by the fact that, although the increase of public debt when private investment is stagnant in depression is necessarily expansionist in its effects (unless, indeed, it undermines public confidence), its repayment out of normal taxation when private investment is already active is not necessarily deflationary; it may, indeed, be just the reverse.[14] Thus the old canon of stability of rates of taxation seems to have little place in a budgetary policy which is designed to control and offset the fluctuations of private business; but the cyclical sensitivity of the various taxes would interact with such a policy at almost every point.

R. F. BRETHERTON

Wadham College
Oxford, England

[14] For some discussion of this point, see Ropke, *Crises and Cycles*, pp. 158–9.

REPORT OF THE CHICAGO MEETING, DECEMBER 28–30, 1936

THE AMERICAN winter meeting of the Econometric Society was held in Chicago on December 28, 29, and 30, 1936, in connection with the meetings of the other social science societies. All sessions were held at the Stevens Hotel.

The meeting opened on Monday morning, December 28, at a joint session with the American Statistical Association and the Institute of Mathematical Statistics, under the chairmanship of Professor Irving Fisher of Yale University. The first paper presented was the Presidential Address by Professor Harold Hotelling of Columbia University on "The General Welfare." Professor Hotelling, in this paper, examined and amended the classical theorem that free competition leads to a maximum of what may be called a component of the general welfare. Two independent components of the general welfare of a strictly economic character may be identified. One of these has to do with the distribution of wealth and income among different persons and classes; the other is the total of the goods and services constituting the national dividend. In addition there are other components of a character not strictly economic, such as personal freedom, civil liberties, good will, artistic appreciation, etc. Considering only the component concerned with the total national dividend, this paper presents as a modernized variant of the classical argument, taking full account of the relations among commodities in a manner impossible to non-mathematical classical theory, the theorem that the social optimum requires that sales be at *marginal* cost, neither more nor less. This implies that excise, sales, and customs taxes are less efficient economically than income, inheritance, and land taxes, even without taking account of the higher cost of collection of the former group. The same theorem implies also what may seem a rather revolutionary proposition, that industries with high fixed costs, such as railroads and electric power plants, might well be subsidized to the extent of their fixed costs, on condition that only marginal costs be charged for their services and products. The argument here is similar to that showing that free bridges are socially more economical than toll bridges.

Professor E. J. Working of the University of Illinois gave a paper on "Some New Indexes of Agricultural Supplies and Carryover."

The final paper of the session was by Dr. Charles F. Roos, Cowles Commission for Research in Economics, on "New Indexes of Stock Prices and Yields and their Relation to the Theories of Capital and Savings." Dr. Roos stated that an examination of the various indexes

of prewar stock-price movements showed that most of them were based on small samples and did not make proper adjustments for changes in capital structure. To meet the need for a reliable history of stock price movements, the Cowles Commission has computed and will publish in 1937 extensive monthly indexes of (1) stock prices, (2) stock prices adjusted for reinvestment of dividends, (3) dividend payments, and (4) yields; and annual indexes of (5) earnings and (6) earnings-price ratios. All proper adjustments have been made for stock rights, stock dividends, and other changes in capital structure. These indexes have been computed for 69 industry groups. They include all common stocks in which transactions on the New York Stock Exchange have been reported during the years 1871–1917; for subsequent years they are based on the Standard Statistics indexes which include over 90 per cent of the total value of common stocks listed on the New York Stock Exchange as well as other leading issues. A comparison of the Cowles Commission indexes with other commonly used ones shows striking differences in location of maxima and minima and in general movements.

The second session, on Monday afternoon, was devoted to Price Analyses, under the chairmanship of Professor Henry Schultz of the University of Chicago. The first paper, by Professor Holbrook Working of Stanford University, was on "The Elasticities of Demand for Wheat." Professor Working pointed out that three calculations of the demand for wheat have received considerable attention. These may be summarized in terms of the coefficients of elasticity of demand stated by the authors in their calculations or implicit in their conclusions:

Gregory King (1699)	$e = -.36$,
R. A. Lehfeldt (1914)	$e = -.$[illegible]1,
G. F. Warren and F. A. Pearson (1928)	$e = -.71$ to -1.23.

The coefficients based on the data of Warren and Pearson measure elasticity of the same "world" "market" demand, but in terms of prices on Kansas farms and in the Berlin market respectively. The difference between the coefficients is almost exactly that to be expected merely from the difference between the average levels of price at these two points within the world market. It reflects a weakness of the coefficient of elasticity as a parameter for summarizing the characteristics of demand. All the foregoing coefficients reflect elasticity of a composite demand which includes demand for storage. Elasticity of demand for actual consumption is so small that even after years of effort devoted to refining the data, no trustworthy measurement has been obtainable. It is possible, however, to set upper limits with a precision that renders them useful:

For food in the United States................	$-e<.008$,
For food and feed in the United Kingdom and the Netherlands...........................	$-e<.07$,
For food and feed in the "world".............	$-e<.10$.

The actual elasticities may be only small fractions of the upper limits given above. It is such elasticities of consumption demand as these which are pertinent to the economic theory of normal price. For market-price theory, coefficients reflecting a composite demand including demand for storage are applicable in principle; but in fact may be highly misleading. They rest on seriously inadequate assumptions with respect to the characteristics of demand for storage.

The second paper was by Professor Lewis A. Maverick of the University of California at Los Angeles, on "Reserve Prices, An Aspect of Dynamic Economics." Professor Maverick summarized his remarks as follows: "The Reserve Price below which the seller will not consent to sell and above which the buyer will not consent to buy is a phenomenon of the business cycle. Reserve prices, expressed in goods rather than in money, were just as real as an exhibition of the stubbornness of the groups of traders in the days before money as they are today. If the group is closely organized, we have the phenomena of monopoly. But even without organization there can be a concert of opinion. When the reserve prices of the two groups differ then there may be either of two results. If that of the buyers is higher than that of the sellers, we have business activity which may lead to recovery and prosperity. If that of the buyers is lower, we have business stagnation and depression. In the latter case it appears that we have overproduction. The period of overproduction will end when one or both of the opposing groups has made concessions from its reserve price so that trading can be resumed. There is a presumption of recurrence in the periods of overproduction even in the barter economy, since the groups will be under continuing temptation to attempt a stubborn policy and since both groups will be practically innocent of any understanding of what happened to them or their fathers in the last economic struggle. The introduction of money alters the recurring phenomena in several ways: it enlarges the market and increases the number of points at which purchase and sale take place. Money as a storehouse of value also introduces the phenomenon of hoarding. Credit also increases the violence and probably the frequency of these phenomena."

The final paper of the session was by Dr. William R. Pabst of Amherst College, on "Butter and Oleomargarine: a Study of Competing Commodities." Dr. Pabst suggested that the problems and interactions of economic phenomena which can be investigated by a consideration of many related commodities rather than independent ones are

worthy of considerable attention. In this study the prices of butter and of oleomargarine, the quantities of these commodities consumed in the United States, income per capita, and time are taken to be interrelated variables and their relationships are evaluated statistically. The equations of relationship are in the form of statistical demand and supply curves in which prices and quantities are alternately chosen as the independent variables. The many possible combinations of these variables provide a fruitful field for further research. The recent developments in the use of indifference-curve analysis are found to be useful tools for the consideration of the equilibrium of two commodities in terms of their prices and quantities and for relating the influence of income. Since income is akin to the cyclical forces which affect the relationships, the comparison of the total amount of money spent upon a commodity with total income appears to have possible ramifications in the investigation of cyclical phenomena. The interrelations of butter and oleomargarine are of practical interest in connection with the present demands for further protection of the butter industry through additional taxation of oleomargarine.

Later Monday afternoon a round-table discussion was conducted on "Uses and Limitations of Mathematical Analysis in Developing Economic Theory," with Dr. Charles F. Roos in the chair.

Professor Harold Hotelling of Columbia University pointed out that the existence of possible systems of indifference curves has been recognized for which the equations of general equilibrium have no solution, or have multiple solutions. In cases of multiple solutions it is important to recognize the difference between stable and unstable solutions. Necessary conditions for stability are that, for each person, the utility shall at the equilibrium point be a maximum subject to his budgetary constraint—not a minimum or minimax. These conditions have been used by the speaker to derive conditions on demand functions (ECONOMETRICA, January, 1935). But it would be interesting and valuable to have some examples of fairly simple utility functions for which the equations of general equilibrium have multiple solutions. For one thing, the possibility of a plurality of *stable* solutions needs to be verified. For another thing, there is need of a more detailed study of the conditions giving rise to Giffen's phenomenon—that is, an increase in the consumption of a commodity such as bread resulting from an increase in its price, on account of the poverty of users who are prevented by its high price from buying expensive substitutes. It may be suspected that where Giffen's phenomenon exists there is another possible solution of the equations of equilibrium, in the neighborhood of which Giffen's phenomenon does not exist. This leads to applications of the theory of maxima and minima in the large.

Professor Eugen Altschul of the University of Minnesota stated that economics was from the beginning an outspoken quantitative science. Logically the Lausanne School has only drawn conclusions implicitly present in the interdependent system of the Classical School. Mathematical analysis becomes even more necessary when the rigid assumptions are dropped and uniquely determined relations are replaced by average relations, and when statistical analysis is used to aid theory. Professor Altschul doubted, however, whether specific dynamic problems, in so far as they included more than variations in time, could be successfully analyzed by mathematics. He thought that it was possible to describe mathematically the relationships but not the actual developments, when movements away from equilibrium created qualitatively new arrangements of economic forces.

Dr. Gerhard Tintner of the Cowles Commission for Research in Economics suggested that one of the most important uses of the mathematical method in economic theory was the possibility of replacing old economic concepts which had some kind of metaphysical connotations. He mentioned especially the work of Hicks and Allen, who (in *Economica,* 1934) discarded utility theory entirely and reduced it to a system of indifference lines or rather indifference complexes in n-dimensional space. They also used the realistic and empirical concept of the elasticity of substitution instead of the old concept of marginal utility. He also referred to the remarkable work of a young Viennese mathematician, Dr. A. Wald, who has shown the conditions under which the Walrasian equations have one and only one solution. His proof (published in *Ergebnisse eines mathematischen Kolloquiums,* edited by Karl Menger, 1935 and 1936), assumes, however, that the utilities derived from different commodities are independent. There is not yet a solution of the general case.

Professor Harold T. Davis of Indiana University commented upon the historical phases of the application of mathematics to a physical or social science. Thus in astronomy the empirical investigations of Tycho Brahe on the motion of the planets was followed by the curve-fitting experiments of Kepler and finally by the a priori formulation of a general mathematical theory by Newton and his successors. In the same manner we may expect these three stages to occur in the historical development of econometrics. The first stage will be the collection of adequate data, the second stage the discovery of exact laws by statistical investigations, and the third stage the formulation of a general mathematical system. We shall probably differ widely in our opinions as to which stage the application of mathematics to economics has now reached.

Emil Lederer of New School for Social Research emphasized one question, already put by Dr. Altschul: There is no doubt that mathe-

matical formulas can be used for symbolizing laws within static economics; and that the main relations within such a static system can be reduced to "laws" which lend themselves to mathematical formulation, if we represent the economic system in its market aspect. The question is as to dynamics. Certainly we can simplify the dynamic system also; but if we do it and if we can, then, express some of its "laws" in mathematical symbols, shall we not, by doing so, eliminate the problems at which we have to work? A harmoniously progressing dynamics could be—in some of its aspects—symbolized in mathematical formulas; but if we make the next step towards reality, inserting one of the elements working for typical and recurrent irregularities—what then? Secondly: is it helpful to represent a development, which in itself will be indeterminate and which therefore will imply alternative solutions, in a formula which will give only one of the final results, that is, one of the equilibriums, that are possible?

The session of Tuesday morning, December 29, was devoted to "Applications of the Theory of Probability to Economic Problems," with Professor Harold T. Davis of Indiana University in the chair.

The first paper was on "Expectations and the Statistical Theory of Errors," by Gerhard Tintner of the Cowles Commission for Research in Economics. The object of Dr. Tintner's paper was to find a general connection between the economic theory of expectations and the statistical theory of errors. The starting point is a utility function which is defined as depending on "arrangements" of a factor x in the course of time, where x may be any economic phenomenon like price, demand, costs, etc. Denoting by $x(i, j)$ the value of the factor at the point in time j, as *expected* at the point in time i, it is easy to derive the maximum conditions and the equilibrium values $\bar{x}(0, j)$ which will yield the maximum satisfaction. If, however, a small error is committed at the point in time h, and the value of x is not $\bar{x}(0, h)$ which would give maximum utility but $\bar{x}(0, h)+e_h$, where e_h is a small error, the situation for the subsequent points in time is different. It is, however, possible to derive a new set of quantities $x(h, j)$ which will, in general, be different from the original $\bar{x}(0, j)$ and will depend on the error e_h which has been committed at the point h. The same general method holds true if new errors are committed at subsequent points k, m, etc. If all these errors are small, then the difference between the original $\bar{x}$ and the maximum values of x, after the errors have been committed, will be linear combinations of the errors. The loss of utility, however, will not be a linear but a quadratic form in the errors. Without any extensive empirical studies it seems very difficult to make assumptions about the distribution of the errors and the correlation between the subsequent errors. The argument is easily extended to the case of a continuous x, in

which case the utility becomes a functional of x. The linear equations, which in the discontinuous case yield the corrections for x, after the error has been committed, are replaced by integral equations.

The second paper was by Messrs. Alfred Cowles 3rd and Herbert E. Jones of the Cowles Commission for Research in Economics on "Some a posteriori Probability Considerations of Stock Market Action." The authors found that evidence of structure in common stock prices was disclosed by computing the ratio of sequences to reversals. A sequence occurs when a rise follows a rise or a decline a decline, and a reversal when a rise follows a decline or vice versa. The method used is more susceptible of interpretation in terms of probability than the techniques frequently resorted to. Units of time considered were 20 minutes, 1 hour, 1 day, 1, 2, and 3 weeks, 1, 2, 3, $\cdots$, 11 months, and 1, 2, 3, $\cdots$, 10 years. For all units of time from 20 minutes to 6 months, with one exception, there was a significant excess of sequences over reversals, indicating a definite element of inertia in stock prices. The evidence of structure suggests the possibility of utilizing this knowledge in a practical way in forecasting. The success of such forecasts depends not only on the ratio of sequences to reversals but also on the brokerage costs and the average change in stock prices during the interval considered. An extensive study, therefore, was made to compute the net gain to be expected from a practical application of this type of forecasting. The largest net gain apparently would occur when the forecasts were based on monthly data, the expected annual net profit then being about 9 per cent. This was based, however, on long-time averages. An investigation of the frequency distributions of the different factors gave the probability of obtaining various net gains or losses. The authors concluded that the ratio of sequences to reversals in stock prices manifests significant structure, but this structure is not sufficiently pronounced or stable to provide a very satisfactory basis for successful speculation.

The third paper, by Mr. William G. Madow of Columbia University, was on "Some Topics in the Theory of Curve Fitting including Periodogram Analysis." Periodogram theory, as developed by Schuster, Turner, Walker, and Fisher, has been concerned only with problems arising from the negation of the hypothesis that no period exists. This hypothesis is negatived by showing that a single period is too important for its importance to be accidental. It is possible, however, that situations might occur in which the hypothesis must be rejected because several periods account for more of the variance than is expected if the null hypothesis is true, even though no single period possesses this property. Distributions by which such situations may be tested are derived analytically. They are, naturally, generalizations of those of

R. A. Fisher.[1] The Schuster periodogram is a special case of problems in analysis of variance and in curve fitting. The solution of the general problem for the analysis of variance, which includes the solution of the problem for curve fitting, yields the necessary distributions as sums of incomplete Dirichlet integrals, of which that obtained by R. A. Fisher is the simplest. Among the possible practical uses of the work is the fact that, to some extent, it makes graphs rigorously usable as statistical tools.

The final paper of the session, "On the Concept of Transvariation and the Law of Comparative Judgment," by Professor Corrado Gini, University of Rome, was read by title in the absence of the author. Professor Gini's own abstract of his paper is as follows:

"When two frequency curves overlap, we say that there is *transgressive variation* or *transvariation* between the two groups of quantities. In this case, if we take at random a quantity from one group and a quantity from the other, there is a certain probability that the difference between the two quantities will be contrary in sign to the difference between the respective averages of the two groups. Such a probability has been called *probability of transvariation.* The *intensity of transvariation,* on the other hand, is given by the ratio between: (a) the sum of those differences (taken in their absolute values) which may be established between the quantities of the two groups, which are contrary in sign to that of the difference between the respective averages; and (b) the sum of all the differences (taken in their absolute values) that may be established between the quantities of the two groups. The probability and the intensity of transvariation are the *constants of transvariation.*

"These definitions have been given in an article of mine[2] published in in 1916, of which a reproduction with some additions is in press.[3] In this article the theory of transvariation is developed, some relations between it and the theories of dispersion and correlation are indicated, and the practical importance of its applications is suggested.

"Since then the theory has received further developments and applications by some Italian statisticians. Castellano has especially studied the effect of sampling on the values of the constants of transvariation.[4] Boldrini has made interesting applications of the constants of trans-

[1] *Proc. Roy. Soc. Lon.*, Vol. 125A, 1929, pp. 54–60.

[2] C. Gini, "Il concetto di transvariazione e la sue prime applicazioni," *Giornale degli Economisti e Rivista di Statistica*, Roma, 1916.

[3] C. Gini, "Memorie di metodologia statistica," Roma, Biblioteca del *Metron*, (in press).

[4] V. Castellano, "Sullo scarto quadratico medio della probabilità di transvariazione," *Metron*, Vol. 11, No. 4, 1934.

variation directed to measure the degree of typicity of the secondary sexual characters.[5]

"In my article, I pointed out that the constants of transvariation cannot generally be deduced from a knowledge of the averages and dispersion of the two groups of quantities. As a matter of fact the constants of transvariation depend also upon the form of the two distributions. When, however, this form is known (as, for example, when both the distributions are normal) it is possible to deduce the constants of transvariation from a knowledge of the averages and dispersion of the two groups. The form of the distribution being known, other deductions are similarly possible: (a) from a constant of transvariation and the dispersion of the two groups, it is possible to deduce the difference between the averages; (b) from a constant of transvariation, the difference between the averages, and the ratio between the dispersion of the two groups, it is possible to deduce the dispersion of the two groups; (c) if the dispersion of the two groups is moreover the same, it is also possible to deduce, from a constant of transvariation, the difference between the averages expressed in terms of the dispersion.

"This is precisely the principle on which the law of comparative judgment of Thurstone[6] is based. Further relations of the said law with the theory of transvariation are expounded and illustrated in the paper here briefly summarized."

The Tuesday afternoon session was devoted to "Multiple Correlation of Correlated Variates." Professor Harold Hotelling of Columbia University presided. The first paper was by Dr. Frederick V. Waugh of the United States Department of Agriculture, "On the Determinateness of Partial Regressions." Dr. Waugh recalled that three recent studies by Ragnar Frisch have formulated and developed the notion of

[5] M. Boldrini, "Su alcune differenze sessuali secondarie nelle dimensioni del corpo umano alla nascita ed alle età superiori," *Archivio per l'Antropologia e l'Etnologia*, Vol. 49, 1919; "I cadaveri degli sconosciuti—Ricerche demografiche ed antropologiche sul materiale della Morgue di Roma," *La Scuola Positiva*, 1920; "Gli studi statistici sul sesso: la proporzione dei sessi nelle nascite e i caratteri sessuali secondari," *Rassegna di studi sessuali*, Vol. 1, 1921; "Differenze sessuali nei pesi del corpo e degli organi umani," *Rendiconti della Reale Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali*, Vol. 29, s.5, 2° sem., Fasc. 1–4, Roma, 1920; "Misure interne ed esterne di alcune ossa lunghe nell'uomo e nella donna," *ibid.* Vol. 33, s.5, 2° sem., Fasc. 7–10, Roma, 1924; *Biometria. Problemi della vita della specie e degli individui*, Padova, Cedam, 1927 (pp. 233–53); *Biometria e Antropometria*, (Vol. 3 del *Trattato Elementare di Statistica*, diretto da C. Gini), Giuffrè, Milano, 1934 (pp. 312–313).

[6] L. L. Thurstone, "A law of comparative judgment," *The Psychological Review*, Vol. 34, No. 4, July 1927; "An experimental study of nationality Preferences," *The Journal of General Psychology*, Vol. 1, 1928.

"scatterances" as a test of the reliability of multiple regressions. The scatterance among n variables is

$$R = \begin{vmatrix} 1 & r_{12} & \cdots & r_{1n} \\ r_{21} & 1 & \cdots & r_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ r_{n1} & r_{n2} & \cdots & 1 \end{vmatrix}. \tag{1}$$

If variable number i is selected as dependent we are particularly concerned with the subscatterance among the remaining $n-1$ variables. This subscatterance is the cofactor, R_{ii}. If R_{ii} is significantly different from zero we can say that the scatter—or lack of correlation—in the subset is significant. In this case the reliability of the regression equation can be tested by the usual methods.

This paper indicates that neither errors of sampling nor errors of computation (if proper checks are used) can account for an observed scatterance different from zero if the true scatterance in the universe is rigorously zero. Errors of observation, however, will always tend to raise the observed scatterance above its true value. We need, therefore, to test the observed scatterance to see if the observed value could be explained entirely by errors of observation.

If the size of the errors of observation can be estimated accurately each correlation coefficient in (1) can be corrected as follows:

$$r_{ij}' = \frac{m_{ij}}{\sqrt{m_{ii} - m_{ii}''}\sqrt{m_{jj} - m_{jj}''}} \tag{2}$$

where m_{ij} is the observed product moment and where m_{ii}'' is the estimated standard deviation of the error of observation for the ith variable. The corrected scatterance and corrected subscatterances can be obtained by inserting the corrected correlations in (1), and calculating a corrected R' and R_{ii}'. If, when generous allowance is made for errors of observation, $R_{ii}'>0$, the usual measures of reliability can be applied to the multiple regression of variable number i on the other variables.

The second paper was "Confluence Analysis and the Standard-Error-of-Parameters Approach to the Problem of Demand for a Producers' Good," by Dr. Francis McIntyre of Stanford University. Dr. McIntyre summarized his paper as follows:

"It has long been recognized by students of statistics that little reliance may be placed upon the normality of the distribution of the ratio to its estimated standard error of a regression coefficient determined from time series. The Student distribution, with its assumption of normality of the parent population, is of little assistance here, although Student's ratio is frequently calculated, 'as the best objective test

available.' Ragnar Frisch has challenged this problem in his recent 'Confluence Analysis.' He abandons standard errors entirely in favor of determinantal functions of the simple intercorrelations of the several variables. These he calls 'scatterances.'

"A statistical analysis of factors affecting the price of refined copper in the United States has furnished an opportunity for the application of the scatterance technique. Space limitations prohibit even a summary of the results, but the tentative conclusions may be generalized: The Frisch methods compel re-examination of the economic content of rejected variables, thus improving the quality of the ingredients fed into the correlation hopper. Further, they provide a test of the applicability to the data at hand of various statistical procedures, assisting for example in the decision whether to use a price index as a deflator or as an explicit variable. So far as this study has proceeded, the economic a priori hypotheses have been much better supported by the scatterance results than by the Student ratios. Relatively simple modifications of the familiar Doolittle solution of the normal equations will yield many of the scatterances Frisch requires, with much less effort than is involved in the method he suggests for their calculation."

The final paper of the session was by Dr. Milton Friedman of the National Resources Committee on "Testing the Significance of the Differences among a Group of Regression Equations." This paper applied the analysis of variance to the problem. N observations on two variables are assumed available. These observations are classified into k groups. A linear regression is computed for each group; it is the significance of the differences among these which is to be tested. The first step is to compute a general regression based on all of the observations. The total sum of the squared deviations about this general regression, S^2, is divided into a part attributable to the variation about the sample regressions and a part attributable to the variation between the regressions. This latter sum of squares is then further subdivided into two portions attributable, respectively, to differences in heights and to differences in slopes. This subdivision is accomplished through using the slope, B', of a regression fitted to the deviations of the observations from their group (or sample) means. The procedure for dividing the total variation into its component parts is summarized in the following table. (In this table the sum of squares about the line with slope B' is designated S'^2; the corresponding sum for the regression for the ith sample, S_i^2; and summation over all samples by $\sum$.) From the various sums of squares, independent estimates of the variance of an individual observation are derived by dividing by the appropriate number of degrees of freedom. The test of the significance of the differences between regressions, between heights, or between slopes

Source of Variation	Degrees of Freedom	Sum of Squares
Total about general regression	$N-2$	S^2
Between sample regressions	$2k-2$	$S^2-\sum S_i^2$
Between slopes	$k-1$	$S'^2-\sum S_i^2$
Between heights	$k-1$	$S^2-S'^2$
About sample regressions	$N-2k$	$\sum S_i^2$

is then made by comparing the relevant variances with the variance about the sample regressions. Whether differences are significant can be determined from the tables of R. A. Fisher or by G. W. Snedecor.[7]

On Wednesday morning, December 30, a session was devoted to "Income," with Dr. Charles F. Roos of the Cowles Commission presiding. The first paper was by Dr. Tord Palander of Stockholm on "The Definition of Production Functions for an Enterprise." A production function expresses the amount of product as a function of the quantities of the factors of production, which are usually assumed to be individually variable and substitutable for each other. Dr. Palander stressed the practical importance of cases where (1) both "limitational" and "substitutable" factors are used and where (2) the entrepreneurial unit comprises several "technical units." A limitational factor can be either "product limitational," a function of the quantity of product, or "factor limitational," a function of the quantity of one (or several) factors used. One cannot speak of the marginal product of one substitutable factor when limitational factors are also needed in the production, because then every change in output is accompanied by a simultaneous variation of the quantities of *several* factors. In the case of enterprises comprising several technical units, a mathematical description of the conditions of production in each unit will require several production functions when there are limitational factors involved. If the units are properly defined, the functions can be determined empirically. The optimal choice of factors to be varied in order to bring about a certain variation of output depends both on the period available for planning and on the period during which the new level of output is to be maintained. The methods propounded in the paper serve to clear the way for more realistic dynamic studies of the output policies of enterprises.

The second paper was by Dr. Alexander Sachs of the Lehman Cor-

[7] After the presentation of the paper, it was pointed out by Professor S. S. Wilks that the same problem had been considered by B. L. Welch in an article entitled "Some Problems in the Analysis of Regression Among k Samples of Two Variables," *Biometrika*, Vol. 27, 1935, pp. 145–60.

poration on "National Wealth, Income, and Debt, and Measures of Financial Soundness and Liquidity."

The third paper was by Professor Irving Fisher of Yale University, on "Income in Theory and Income Taxation in Practice." Income had been defined by the speaker in *The Nature of Capital and Income* in 1906 as services received. The bulk, in value, of services received is in money—dividends from stocks, interest from bonds, etc. In the present paper the subject is reviewed and extended by the use of mathematics. One object is to show the fallacies underlying our present income tax. Our income-tax laws make capital levies masquerade as income taxes. They therefore lend themselves to tax evasion and avoidance. We may reduce the problem of approximating the income of an individual to measuring the use or consumption of wealth—"real income"—in terms of money. To do this we assume that most of the consumption services are equivalent to the money spent for them. The only exceptions worth attention are three: house shelter, house furnishings, and automobile. The problem therefore reduces to calculating these three discrepancies, each between a year's use and its cost. If Smith buys a new automobile he pays more than the value of one year's use and the cost-use discrepancy is negative. If he lives in his own house, fully paid for before the current year, the cost-use discrepancy for housing is positive. The tax schedule for an income tax thus calculated for "real income" will differ from the present tax schedule in several important particulars.

The final session was held on Wednesday afternoon, under the chairmanship of Professor Harold Hotelling, Columbia University. The first paper, "Chinese Influences upon the Physiocrats," was by Professor Lewis A. Maverick of the University of California at Los Angeles. Dr. Maverick spoke of the effects on French thought of reports from a group of highly intellectual Jesuit missionaries sent to China under royal auspices during the period from 1685 to 1773. These reports reflected a sincere admiration for the civilization of China, and in addition they were edited by the Jesuits at home to influence the religious and political situation in Europe. Quesnay, dissatisfied with the treatment of agriculture in France, found glowing reports of the elevated position of agriculture in China. Seeking a government by individuals less willful than his master, Louis XV, he read that China was governed by a group of mandarins selected by a rigorous process of successive examinations and daring even to criticize the emperor when he departed from the ancient and just laws of that empire. Seeking for a means whereby the divinely ordained natural order might be made effective, he wrote approximately one hundred pages in the *Éphémérides* on the subject of the Despotism of China, in which he

developed his theory of "legal despotism." European scientists also considered the question of reconciling ancient population figures of China with the possibilities of world population growth after the traditional date of Noah's flood. In general, the effect of China upon philosophical and economic thinking in Europe paralleled its effects upon the arts.

The second paper was by Dr. August Loesch of the University of Bonn, on "Population Cycles the Cause of Business Cycles." Dr. Loesch showed that there are large cycles in the increase of the German population as a whole, and in the increase of labour supply in particular. These fluctuations are not very much affected by economic events, their main cause being the great wars. Hence they may unhesitatingly be compared as an independent factor with the fluctuations of business activity. This comparison, which covers the period since 1860, shows a pretty close co-variation of the variations in the rate of increase of population and of production. The theoretical explanation of this fact is as follows: an increase in labour supply increases the demand for capital goods, especially for houses, factories, and machinery. This demand is more reliable and more important in value than the one depending upon technical innovations and is therefore more likely to bring about an upswing. Once this increase has continued for some successive years, and the industry of capital goods has adapted itself to it, its unexpected diminution leads to overproduction, which means depression. In this way most of the German business cycles can be explained.

The final paper was "A New Index of Unemployment," by Dr. Alexander Sachs of the Lehman Corporation.

Dickson H. Leavens

Cowles Commission for
Research in Economics
Colorado Springs, Colorado

ATTENDANCE AT THE OXFORD MEETING, SEPTEMBER 25–29, 1936

THE SIXTH European meeting of the Econometric Society was held at New College, Oxford, England, from September 25 to 29, 1936. A report of the proceedings will appear later in ECONOMETRICA. The following persons attended: Johan Åkerman, Lund; R. G. D. Allen, London; Harold Barger, London; A. S. J. Baster, Exeter; Dr. H. Bolza, Würzburg; Miss M. Bowley, Oxford; R. F. Bretherton, Oxford; A. J. Brown, Oxford; E. H. Phelps Brown, Oxford; Dr. F. Burchardt, Oxford; Dr. A. Bijl, Amsterdam; D. G. Champernowne, Cambridge; W. A. Chudnowski, Philadelphia and Oxford; Colin Clark, Cambridge; Miss Ruth Cohen, Oxford; W. de Langen, Holland; Dr. J. J. J. Dalmulder, Rotterdam; Prof. G. del Vecchio, Bologna; V. Edelberg, London; C. Eisenhart, Princeton and London; M. R. el Shanawany, London; Dr. M. J. Elsas, London; Prof. L. M. Fraser, Aberdeen; Prof. Ragnar Frisch, Oslo; R. Gibrat, Paris; T. Haavelmo, Oslo; R. L. Hall, Oxford; Dr. L. Hamburger, Scheveningen; R. F. Harrod, Oxford; J. R. Hicks, Cambridge; Ursula Hicks, Cambridge; C. J. Hitch, Oxford; Miss M. Joseph, Geneva; Dr. M. Kalecki, Warsaw; Dr. L. Lachman, London; Dr. A. P. Lerner, London; M. Landau, Cambridge; Prof. E. Lindahl, Göteborg; Dr. A. Löwe, Manchester; G. D. A. Macdougall, Leeds; Miss H. Makower, London; Dr. J. Marschak, Oxford; J. E. Meade, Oxford; Dr. H. Mendershausen, Geneva; G. R. Nair, London; U. S. Nair, London; Dr. A. M. Neuman, London; Dr. J. Neyman, London; S. Nikodym, Warsaw; Dr. J. M. I. Reitsma, Holland; H. W. Robinson, London; Paul Rousseaux, Charleroi; Dr. J. Rutski, Vilna; G. L. S. Shackle, London; Dr. Hans Staehle, Geneva; J. Guha Thakurta, London; Dr. J. Tinbergen, Geneva; B. D. White, Oxford; O. Williams, London; K. Williams, London; J. Wisniewski, Warsaw; Dr. H. Zassenhaus, London.

ANNOUNCEMENT OF THE DENVER MEETING, JUNE 21–26, 1937

A MEETING of the Econometric Society will be held at Denver, Colorado, during the week of June 21–26, 1937, in connection with the annual summer meeting of the American Association for the Advancement of Science. The program is being arranged by Dr. Charles F. Roos, 420 Lexington Avenue, New York City.

ANNOUNCEMENT OF THE SEVENTH EUROPEAN MEETING, NEAR GENEVA, SEPTEMBER 11–15, 1937

THE SEVENTH EUROPEAN MEETING of the Econometric Society will be held near Geneva, most likely at Faucille, September 11 to 15, 1937. The date has been chosen somewhat earlier than last year in order to enable some of our American members to attend the meeting. It is planned to organise colloquium lectures on:

(1) Methods of measuring and eliminating changing seasonal components;
(2) Methods of solving linear differential and difference equations;
(3) Multiple Correlation of Time Series.

There will also be reports read on subjects immediately related to Genevese work, such as measurement of terms of trade, of recent developments in the study of food consumption, of interest rates, of index numbers. There will probably also be some papers on other topics. Account will be taken, however, of the desire not to overburden the programme. Even more time than at previous meetings will be left for informal colloquia discussions. These have proved to be a great success. A full announcement will follow in the July issue of ECONOMETRICA.

J. TINBERGEN

REPORT OF THE COUNCIL

THE COUNCIL of the Econometric Society has elected the following officers for 1937: Harold Hotelling, President; Arthur L. Bowley, Vice-President; and Alfred Cowles 3rd, Secretary and Treasurer.

The Council members elected by the Fellows for terms to expire in December, 1939 are Costantino Bresciani-Turroni, Irving Fisher, Charles F. Roos, and Joseph A. Schumpeter. Council members whose terms expire in December, 1937 are Ragnar Frisch, John Maynard Keynes, and F. Zeuthen. Those whose terms expire in December, 1938 are Albert Aupetit, Arthur L. Bowley, and Wladyslaw Zawadzki. Ex officio Council members in 1937 are Harold Hotelling, President of the Society, and Alfred Cowles 3rd, Secretary and Treasurer.

It was voted by the Council to define the term of office of the Editor of Econometrica to be four years, beginning January 1, 1937. Ragnar Frisch was elected Editor for this term.

ALFRED COWLES 3RD
Secretary

OBITUARY

ECONOMETRICA records with deep regret that the deaths of the following members of the Econometric Society have occurred during the past year.

M. Henry Le Chatelier
75 rue Notre Dame des Champs
Paris VI, France

Dr. Dobroslav Krejci
Masarykova Universita
ul. Macku 68
Brno-Brünn, Czechoslovakia

Col. Malcolm C. Rorty
Old Spout Farm
Lusby, Calvert County
Maryland

ANNUAL SURVEY OF ECONOMIC THEORY: THE RECENT CONTROVERSY ON THE THEORY OF CAPITAL

By Nicholas Kaldor

The last few years have witnessed the emergence of a tremendous literature on the theory of capital and interest—stimulated, no doubt, by the urgency of finding the appropriate theoretical criteria for a policy designed to mitigate economic instability. A large part of this literature has been directly concerned with the question how far the concept of the "period of production" is relevant for an analysis of industrial fluctuations. Another part, digging deeper into the problem, dealt with the *prima facie* question how far traditional capital theory, formulated under the hypothesis of a stationary state, still retains its validity in essential features once this hypothesis is abandoned. These writings were mainly concerned with the problems of expectations, foresight, uncertainty. Finally, largely owing to the offensive launched by Professor F. H. Knight, there was a revival of the discussion on the fundamentals of capital theory itself, comparable in nature to the famous controversy between J. B. Clark and Böhm-Bawerk in the first decade of the century. In this controversy the problems introduced by dynamic changes were not so much in question as the legitimacy of the "investment period" theory of capital even within the narrow framework of static assumptions. Professor Knight's attack[1] has been taken up and supported by other writers,[2] has been fre-

[1] The following articles by Professor Knight deal mainly with this question: "Capital Production, Time and the Rate of Return," *Economic Essays in Honour of Gustav Cassel*, London, 1933, pp. 327–342; "Capital, Time and the Interest Rate," *Economica*, August, 1934, p. 257; "Professor Hayek and the Theory of Investment," *Economic Journal*, March, 1935, p. 77; "The Ricardian Theory of Production and Distribution," *The Canadian Journal of Economics and Political Science*, February, 1935; "The Theory of Investment Once More: Mr. Boulding and the Austrians," *Quarterly Journal of Economics*, November, 1935; "The Quantity of Capital and the Rate of Interest, Part I," *Journal of Political Economy*, August, 1936; "Part II," *ibid.*, October, 1936. (The last of these appeared too late for consideration in the present article.)

[2] The following authors could be regarded as supporting Knight's criticism in varying degrees: M. F. Joseph and K. Bode, "Bemerkungen zur Kapital und Zinstheorie," *Zeitschrift für Nationalökonomie*, Vol. 6, June, 1935; H. S. Ellis, "Die Bedeutung der Produktionsperiode für die Krisentheorie," *Zeitschrift für Nationalökonomie*, Vol. 6, 1935; Nurkse, "The Schematic Representation of the Structure of Production," *Review of Economic Studies*, June, 1935.

The following articles, recently published, deal with more or less the same problems though they are not directly related to the issues of the present controversy: C. H. P. Gifford, "The Concept of the Length of the Period of Production," *Economic Journal*, December, 1933, p. 611; "The Period of Production

quently reiterated by Professor Knight himself, and, on the Austrian side, has been answered by Professor F. Machlup and Professor F. A. von Hayek.[3] It is with this particular controversy that the present article will be concerned.

The literature created by this discussion is already sufficient to fill volumes, and most of it makes very difficult and often tedious reading. Yet a perusal of the more recent publications does not suggest that much progress has been made towards mutual understanding. While Professor Knight's position and those of other critics is not entirely acceptable in the view of the present writer, it appears that on the Austrian side none of his chief points have yet been fully understood or effectively answered.

For this state of affairs, I think Professor Knight is partly responsible. A serious reading of his numerous articles on this particular subject does not make it easy to discover the essential points of departure. He makes so many points that one is apt to get lost among them, not knowing how to distinguish between the primary and secondary, the important and the unimportant; while the conclusions are frequently clothed in paradoxical sentences which are intended to challenge the

under Continuous Input and Point Output in an Unprogressive Community," *Econometrica*, Vol. 3, April, 1935, p. 199; K. E. Boulding, "The Theory of a Single Investment," *Quarterly Journal of Economics*, Vol. 49, May, 1935, p. 475; "Time and Investment," *Economica*, May, 1936, p. 196; J. Marschak, "A Note on the Period of Production," *Economic Journal*, Vol. 44, March, 1934, p. 146; J. Marcus Fleming, "The Period of Production and Derived Concepts," *The Review of Economic Studies*, Vol. 3, October, 1935; A. Smithies, "The Austrian Theory of Capital in Relation to Partial Equilibrium Theory," *Quarterly Journal of Economics*, Vol. 50, November, 1935; V. Edelberg, "Elements of Capital Theory, A Note," *Economica*, August, 1936; Karl H. Stephans, "Zur neureren Kapitaltheorie," *Weltwirtschaftliches Archiv*, January, 1935; "Zur Problematik der Zinstheorie," *Zeitschrift für Nationalökonomie*, Vol. 7, 1936; Richard von Strigl, "Zeit und Produktion," *Zeitschrift für Nationalökonomie*, Vol. 6, 1935; E. Schneider, "Das Zeitmoment in der Theorie der Produktion, I," *Jahrbücher für Nationalökonomie und Statistik*, 1935; "II," *ibid.*, 1936; A. Mahr, "Das Zeitmoment in der Theorie des Produktivzinses," *Zeitschrift für Nationalökonomie*, Vol. 7, 1936; Carl Iversen, "Die Probleme des festen Realkapitals," *Zeitschrift für Nationalökonomie*, Vol. 7, 1936; O. Lange, "Interest in the Theory of Production," *Review of Economic Studies*, Vol. 4, June, 1936; H. Gaitskell, "Notes on the Period of Production," *Zeitschrift für Nationalökonomie*, Vol. 7, 1937.

[3] F. Machlup, "Professor Knight and the 'Period of Production'," *Journal of Political Economy*, Vol. 43, October, 1935, p. 577 (together with Professor Knight's comment), and a further Rejoinder, *ibid.*, December, 1935, p. 808; F. A. von Hayek, "The Mythology of Capital," *Quarterly Journal of Economics*, Vol. 50, February, 1936, p. 199. Reference should also be made to another article by Professor Hayek, dealing with earlier criticisms and a further elucidation of his views, "On the Relationship Between Investment and Output," *Economic Journal*, June, 1934, p. 207.

mind but without a sufficient indication of where to turn in order to uncover those mental processes which must have led up to them.

The aim of the present article is to review the essential points in Professor Knight's argument, to examine them in the light of other criticisms which have been put forward, and, finally, to analyse to what extent and in what respects they destroy the validity of traditional theory. Since this reconstruction of Knight's views has involved some "filling in" of gaps in the printed argument at certain stages, it is not necessarily a "correct" version of his views; it should be considered as an interpretation rather than a summary; and it is possible that it will be repudiated by the author himself.

II

Professor Knight's criticism of the "Austrian" doctrine can, I think, be summarized under three headings: First, that it is impossible to distinguish between permanent and nonpermanent resources (or "original" and "produced" means of production) or between the services of those resources; Second, that it is irrelevant, and in many cases, impossible, to distinguish—analytically or physically—between expenditures incurred in "maintaining" resources and those incurred in "replacing" them; Third, that there is no necessary correlation between the "period of production" and the quantity of capital. Among these, perhaps, the second is most open to criticism and at the same time, least important; whereas the third is certainly the most important and at the same time the most inadequately explained. But let us deal with each of these points in turn.

1. *Permanent versus Nonpermanent Resources.*—Here Professor Knight makes use of two separate arguments. In the first place he sharply distinguishes between the services of resources and the resources themselves (the actual physical objects from which the services flow). The former, in his view, cannot be thought of except as a rate of flow in time: like light or electricity (but unlike water) they flow, but cannot exist as a stock, or have their use transferred to any other period. Just as one cannot "bottle up" sunshine—except in the sense of transferring its energy into some other object, like oranges, which means "consuming" it by creating value in that object—today's labour hours cannot be deferred until tomorrow: they must be used immediately or lost. As regards the latter—pieces of land, labourers, and machines—no distinction can be drawn between permanent and nonpermanent resources, simply because permanent resources—apart from a few and insignificant exceptions—do not exist. It is essentially a fiction that there are "permanent" resources which exist without being maintained and whose services are therefore forthcoming at a rate inde-

pendent of their price. This fiction is admissible in static or stationary-state analysis, where it does not affect the immediate issues involved; but it is inadmissible to treat it as a relevant *fact* upon which a theory may be built. That it is fiction and not fact is shown by the reflection that neither land nor labour services would continue to flow (from the same resources) without the application of current services for their maintenance. No type of natural resources truly possesses "indestructible powers"; the best that can be expected is that the flow of services can be kept up permanently by continued maintenance.[4] A piece of land can be kept permanently in good condition by careful husbandry; but its "consumption" (in the same sense that capital goods can be consumed) is certainly possible by reducing its value to nil through nonmaintenance. In fact some types of resources (such as sources of coal and oil) cannot be kept intact however much is spent on their maintenance, though how long they last and the amount of services yielded may be influenced by expenditures on their upkeep.[5]

The point is equally obvious in the case of labour. The services flowing from a labourer could not be forthcoming unless he is given food nor could he be replaced after his death unless children are "maintained" until they reach the age when valuable services begin to flow from them (during their "construction period"). This way of looking at the matter would not sound so ridiculous but for the historical accident of the abolition of slavery. In a slave state, investment in human labour is in all respects identical with investment in machinery. And even in the nonslave state there is a minimum price necessary to maintain the labourer, while the Malthusian theory of reproduction applies, in certain countries and periods, to a considerable extent.

Even if the maintenance of labour does not proceed on strictly economic grounds in a world where everyone owns his own labour—since the preference for life over death cannot be expressed in marginal terms—maintained (and replaced) it must be; and therefore all re-

[4] The most important exception to this is sunshine which—given static weather conditions—flows at even rate without anything being done to the sun. But since neither sunshine nor the sun can be made subject to human property rights and thus market valuation, this exception is irrelevant. It might be argued also that sheer area (involving exposure to light and rainfall and power to support structure) *is* an "original and indestructible power of the soil" in the Ricardian sense and the only one; but even here we must qualify that area may shrink in some cases (e.g., on river banks) without maintenance.

[5] Professor Knight would go further and say that such nonexhaustible resources can also be "maintained" permanently by creating resources whose services provide a substitute for them. This view is justified, only in so far as *perfect substitutes* can be found (which is by no means always the case; not all uses of coal can be equally replaced by water power).

sources (i.e., *all* scarce objects, including human beings) must have some input or maintenance stream in order to have a permanent output stream (both of which are, of course, to some extent variable). No distinction can be drawn along this line; and the criticism urged against Professor Knight,[6] that he regards capital as maintaining itself permanently without maintenance expenditure, misses its point. From one standpoint all resources are "permanent"—which merely implies that, if they are maintained, they are maintained; while from another standpoint, none are permanent—since none will remain unconsumed unless maintained. What matters is that no distinction can be drawn between permanent and nonpermanent resources, whichever standpoint is adopted.[7]

Professor Knight's second argument in this connection refers to the analogous, but by no means identical, distinction between "original" and "produced" factors. Even if the distinction between permanent and nonpermanent resources is invalid, this latter distinction would still be valid, if it were true that the services of one set of resources—the "original factors"—produce another set of resources, the services of which—either by themselves, or with the aid of the services of the former—produce want-satisfying service flows. But there is no such one-way causation as is assumed by the Austrian theory. Resources are produced with the aid of the services of all kinds of resources; and it is even conceivable that the services of produced resources *by themselves alone* and without any aid from the services of "nonproduced" resources, should produce an endless succession of further produced resources. (It is "conceivable," but I think Professor Knight will admit that such an eventuality is not very likely.)

I hope to show later on that the importance of this latter point has been rather exaggerated—at any rate if it still remains true that the services of "produced" resources always require the co-operation of

[6] Cf. Hayek, "The Mythology of Capital," *op. cit.*, p. 214. "The very concept of capital arises out of the fact that, where nonpermanent resources are used in production, provision for replacement of the resources used up in production must be made, if the same income is to be enjoyed continually, and that in consequence part of the gross produce has to be devoted to their production." But are there any resources for which this is not true?

[7] Moreover, even if it were true that some resources are permanent (in the sense of requiring *no* maintenance) whilst others are not, this fact would not really be relevant from the point of view of capital theory. As will be shown below, "permanent resources" might very well be "capital goods," so long as they are augmentable in quantity; while there are various "nonpermanent goods" which are not part of capital (in the sense that they do not enter into the determination of the rate of interest) for the simple reason that their quantity cannot be augmented. In any case, the distinction between permanent and nonpermanent goods cannot be used to demarcate capital from other resources.

the services of "nonproduced" resources in further production. Professor Knight is quite right in insisting, however, that it destroys Böhm-Bawerk's concept of a "period of production." If the services of produced resources become embodied in further resources (and so on, in endless succession) there is no definite time lag between the investment of a "service unit" and the corresponding emergence of another service unit which is instantaneously destroyed by consumption. The "investment period" for certain services invested on a particular date (or, rather, for a small portion of those services) might be infinity. But this does not imply, in our view, that it is impossible to attribute an "average investment period" for the services *embodied* in a given stream of consumption goods.

It might be argued that the services of the resources *accruing at the present moment* might be regarded as "original factors" as against the services of resources accruing at any subsequent moment. Such a distinction, however, would be meaningless when applied to the time continuum of static equilibrium; and it is questionable whether the periods for which the services accruing at a single moment are invested, are in any way definite in the absence of stationary conditions. For the inputs of different dates jointly produce the outputs of different dates; and it is impossible to separate out the contribution to the output of different dates of the input of a single date.[8,9] This is the chief objection against the concept of an "investment period of currently accruing services" (as against the investment period of the services embodied in a given stream of consumption goods) which Professor Hayek now regards as relevant.[10] Another (alternative) objection is

[8] It is only under the assumption of stationary conditions, where both the output stream and the input stream are constant over time, that an investment period can be imputed to the input of a particular date; since in this case, this period will equal the investment period of the services *embodied* in the capital goods. Cf. also p. 212.

[9] This has already been stated by Wicksell, *Lectures on Political Economy*, English edition, Vol. I, London, 1934, p. 260. Wicksell was considering the analogous problem (or, rather, the same problem from the "other end," so to speak) whether the amount of labour disinvested by the "annual use" of a machine can be measured. ". . . fundamentally it is just as absurd to ask how much labour is invested in either one or the other annual use as to try to find out what part of the pasture goes into wool and what part into mutton. It is only at the margin of production that these quantities can be differentiated and have a concrete significance attached to them." Assuming variability at the margin, it is possible of course to determine by how much the output of various dates can be increased by a *marginal increment* of the input of a single date. But this does not imply, as Machlup appears to believe ("Professor Knight and the 'Period of Production'," *op. cit.*, p. 587) that it is possible to evaluate the contribution of the input of a given period to the output of different future periods.

[10] "The Mythology of Capital," *op. cit.*, pp. 206, 218–219.

that, in the absence of stationary conditions, this measure would be correlated with changes in the scale of new investment, rather than changes in the quantity of capital. It might easily remain constant while the quantity of capital is increasing if accumulation proceeds at a steady rate; while it could actually diminish if the rate of accumulation slowed down.

2. *Maintenance versus Replacement.*—Professor Knight argues in the second place that the maintenance expenditure (which we have seen is necessary for all resources) cannot be distinguished from expenditure incurred to replace worn-out capital goods. The usual distinction between replacement and maintenance is based on the idea that the former does (while the latter does not) bear a definite ratio to the service life of particular capital goods. This is best elucidated by an example. If the investment in a particular stock of houses is not maintained—the amortization funds are not put aside year by year—the amount thus "released" will bear a mathematical relation to the service life of the houses (a relation varying with the rate of interest, but definite at any given rate). If, on the other hand, "maintenance expenditure" in the narrower sense is not incurred (the roof leakages are not stopped, etc.) the house may become immediately useless and the destruction in value caused thereby bears no relation to the amount "released." Now, in the case of many capital goods no definite "replacement" ever occurs; the maintenance may consist only in the periodic replacement of "individual bits"; but that type of replacement need bear no relation to the shortening of service life (or, rather, the reduction in the discounted value of future services) caused by a reduction in maintenance expenditure. A railway locomotive, for example—apart from changes in knowledge, causing technical obsolescence—is never entirely replaced although every single part of it might be exchanged in the course of time, as this becomes necessary. But the sum of such maintenance expenditures cannot be brought into any simple relation with the cost of the locomotive as a whole; and failure to incur such expenditure in any particular respect (e.g., the replacement of a piston) will not destroy *part* of the value of the locomotive; it will destroy its entire value.[11]

[11] This, I believe, is also the reason for the view, which most people found so puzzling, that the "investment period" of the services of resources must be either zero or infinity; i.e., zero for the services engaged in producing current output-streams (from existing capacity) and infinity for the services employed in creating new "capacity." It does not imply a denial that capacity requires maintenance, but merely the view that no definite investment period can be attributed to the services employed in such maintenance for the simple reason that such expenditure is the absolute condition of the "functioning" of the capacity rather than the cause of a definite prolongation of its service life. In the above

Moreover, if "replacement" occurs regularly and continuously—and we shall see presently Professor Knight's reasons for regarding it as if it did—"replacement expenditure" becomes undistinguishable from "maintenance expenditure" in the narrower sense; and, therefore, according to Professor Knight, *the two should be lumped together, and not treated separately.*[12] I am not sure that even so, with a little mental effort, it would not be possible to forge a criterion for an analytical distinction, but I certainly do not think it would be worth the trouble. As we shall see later on, the essential point of the "Austrian" theory of capital does not really depend upon the validity of this distinction.

3. *The Optimal Length of the Investment Period.*—None of the points mentioned so far affect the fundamental assumption of the Austrian theory: the law of roundaboutness. Now we come to the argument with which Professor Knight seeks to prove that this law, irrespective of whether it is true in reality or not, is irrelevant from the point of view of capital theory, for it cannot be shown that an increase in the quantity of capital in a community will necessarily imply the adoption of more "roundabout" processes.[13] In order to show that this argument is independent of the previous objections, we shall assume for the present that "maintenance" does consist of periodic replacement of capital goods, as the Austrian theory apparently assumes, and that capital goods are exclusively produced by the services of other resources, i.e., labour. Let us revert therefore to the traditional situation exemplified by a world where only houses are produced and only labour is required to build (or replace) such houses. The only consumption good will then be the services flowing from houses, i.e., "room-years"; and we might assume the coexistence of different types of room-years. We shall defer for a moment the question how the "degree of round-

case of the locomotive, the labour engaged in building it remains invested for an infinite period, if the locomotive is kept in repair, but only for a very short period—perhaps a day—if the necessary repairs are not made good. Similarly with the labour engaged in making repairs. It is impossible to say by how much the service life of a locomotive is prolonged by the replacement of a worn-out piston. If it is not replaced, the future service life of the locomotive becomes zero, while if it is (and all other "pistons" are also replaced in the course of time) its lifetime might be infinity.

[12] Cf. especially "The Theory of Investment Once More: Mr. Boulding and the Austrians," *op. cit.*, p. 59: "the process of amortization and replacement is precisely the continuance of an old life and not a new birth"; also "particularly with reference to increments of value, capital as capital, it seems truistical to say that if it is kept in existence there is no amortization and replacement but only continuous maintenance."

[13] I am indebted to Mr. Milton Friedman, of the National Resources Committee, Washington, for helping me to understand Knight's argument in this connection.

aboutness" is to be measured; under these assumptions it will obviously vary with the lifetime of the houses. The famous Jevons-Böhm-Bawerkian law is satisfied if we assume that for each particular type of house (i.e., a type of house is one which provides a given kind of room-year) it is always possible to increase service life in a given proportion by increasing the construction costs of the houses in a lesser proportion.

We shall make two further assumptions, which, in my view, are also implicit in Knight's analysis. The first is that there is perfect competition and constant returns to scale (i.e., the production function is homogeneous in the first degree). The second is that investors have static foresight regarding the future, which implies that they expect the continuance of the same prices in the future as are ruling at present.

Under these assumptions the "optimum degree of durability," i.e., the optimum length of service life of houses, will be the one which maximises the rate of return on a given quantity of investment. In case of resources, such as houses, which are assumed to be periodically replaced, it is not immediately clear how this rate of return is obtained. It will obviously depend on the building cost of the houses (on the price of labour) and on the price of room-years; but it will also depend on the way amortization is provided. The representative investor, in deciding upon the degree of durability he should adopt, will deduct from the expected annual (gross) income of the house a sum sufficient for its replacement when it falls due. The net return of the investment will thus depend on the annual amount of this deduction, i.e., the annual amortization quota. It is only when the relative costs of amortization of the different types of houses are known that it is possible to determine the optimal length of service life.

But the amount of this annual deduction, given the length of service life, will obviously depend on the rate of interest at which the amortization quotas are accumulated. The higher this rate the lower the annual sum required to secure a given "replacement fund" at the end of a definite period; and the higher, in consequence, the rate of return on the investment itself. Now the rate of interest at which the amortization quotas are accumulated can certainly not be higher than the rate of return on the investment, since this would imply the existence of an investment opportunity which is superior to the one in question, in which case that particular investment would never be adopted. For similar reasons, it cannot be lower than the rate of return, since this would imply that the amortization quotas are invested in an investment opportunity which is inferior to the one in question, and the investor always has the choice of reinvesting his capital in the same uses in which it was originally invested. *Consequently the two must be*

equal to one another: and this condition makes the rate of return on a particular type of investment uniquely determinate. *The real rate of return on a particular type of investment is therefore that rate which satisfies the condition that the rate at which the amortization quotas are accumulated is identical with the yield of the investment itself.*[14] The optimum "degree of durability" is the one which maximises the rate of return, calculated in this manner.[15]

This can be elucidated by the following example. Let us assume that the same house (i.e., a house having the same number of rooms, of exactly the same type) can be built in three different degrees of durability. The first costs 1000 units to build and lasts thirty years. The second costs 1100 units and lasts forty years. The third 1200 units and lasts fifty years. We shall calculate first the net yield of the three houses by assuming that the amortization quotas are accumulated at various "given" rates of interest, and second, we shall calculate the real rate of return for each type by assuming that the amortization quotas are accumulated at the same rate as the "net yield" itself. The following table shows the comparative rates of return under the two assumptions:—

Rates of Interest Used in Calculating Amortization	Net Yield (per cent) of		
	Type I	Type II	Type III
per cent			
2	4.8	5.0	5.0
3	5.4	5.5	5.4
4	5.7	5.8	5.5
5	6.0	6.0	5.7
6	6.2	6.2	5.8
(7	6.4	6.3	6.0)*
Real Rate of Return	6.35	6.2	5.9

* At seven per cent none of the investments would be undertaken, since none would have a yield equal to that rate.

[14] The real rate of return, as defined above, is necessarily the same as the one which equates the sum of the discounted *gross* returns of a house (with no deduction for amortization) with its costs of reproduction. It is identical therefore with Professor Fisher's "rate of return over cost" (*The Theory of Interest*, pp. 155 ff.), Wicksell's "real" or "natural" rate, and the "internal rate of return" of Mr. Boulding. ("The Theory of the Single Investment," *op. cit.*, p. 479.) But it is only under the assumption of constant (value) returns to scale (from the point of view of the individual investor) that the optimal mode of investment can be determined by the condition that the real rate of return is maximised. Under conditions of diminishing returns to scale the determination of the optimal method of investment is more complicated and presupposes that the rate of interest is already known.

[15] This conclusion is true, irrespective of whether the output or input streams

It is easily seen that for each type of house the net yield will be at its maximum at the "real rate of return." This is the return which the investment yields if the amortization quotas are reinvested in the same use as the one represented by the original investment. This in turn implies that the investment—after a certain lapse of time, at any rate—is so arranged that the amount of capital invested in a given use is kept at an (approximately) steady and even level over time; this means, in real terms, that the age distribution of houses of each type remains constant in successive periods of time. If individual houses last, e.g., 30 years, a "house investment" will consist of a series of 30 houses, varying in age between 0 and 29 years, one of which is replaced every year. The gains from the investment of a certain amount of capital are therefore only maximised if the time quantity of the investment is stabilised: unless it pays to do the latter it does not pay to undertake the investment at all. Such a "staggering" of capital is thus an indispensable condition of a state of equilibrium.[16]

There need be no difficulty in arranging a maintenance scheme of this type, at any rate under the idealized conditions assumed in the theory. "Houses" may be big; (too big for the individual investor to buy a series of 30 houses), but, if not houses, at any rate the ownership titles in those houses are divisible: and so it ought to be possible for anybody to arrange his investment in such a way as to keep the amount of the investment per unit of time constant. To achieve this end may be considered, therefore, as one of the functions of the capital market.[17] All that is necessary to assume is that the indivisibilities do not go so far as to prevent the coexistence of a sufficient number of houses of each type and age.

This is the meaning of Professor Knight's repeated assertions that capital goods ought to be treated as if they were permanently and continually maintained, that capital is perpetual or a "permanent

---

are uniform over time (as assumed in the text) or not. Whatever the time shape of output and input streams, there is only one rate of interest, corresponding to any given constellation of outputs and inputs, which makes the discounted value of all outputs minus the sum of the discounted values of all inputs (including the initial input, or construction cost), for any given date, equal to zero. And since all possible constellations of the time shapes of output and input streams are given by the production function, there will be (normally) only one possible time constellation of inputs and outputs which makes this "internal" rate a maximum. Cf. also Knight, "The Quantity of Capital and the Rate of Interest, Part I," *op. cit.*, p. 445.

[16] This has been stated by Wicksell and set out at length by Åkerman, *Realkapital und Kapitalzins*. Cf. also Wicksell, *Real Capital and Interest*, Lectures, I, pp. 258 ff.

[17] Moreover, it is sufficient to assume that this is possible for *some* investors, since they, through the workings of competition, can prevent the others from investing anything at all in that particular type of investment.

fund." Investing in 30 houses, one of which falls due for replacement and is *planned to be replaced* every year *ad infinitum*, is the same thing as investing in a house which lasts forever, while a certain sum has to be paid out every year to keep it in repair. This sum can be looked upon as "maintenance cost"; it can also be looked upon as the contribution of the services of other resources needed to produce the room-year service which is instantaneously consumed.[18] Thus every investment should be regarded as the source of a certain output stream and the consumer of a certain input stream (both of which are, of course, to some extent variable), in addition to which it will have a certain "initial input" or construction cost. As Professor Knight has shown, in the case where these streams are constant over time, the relation of output value to input value determines the investment period (in his terminology, the turnover period[19]).[20] Since the annual net income

[18] This is also the meaning, I believe, underlying Knight's statements that "maintenance is merely a detail of administration," or that "capital is an integrated, organic conception." What it means is that, in a state of equilibrium, all capital, however durable or perishable are the individual capital goods of which it consists, must be regarded as a fund which is continuously maintained—it cannot be thought of otherwise—since its yield can only be maximised on this basis.

[19] "The Theory of Investment Once More," *op. cit.*, p. 55. According to Professor Knight, this turnover period has only meaning "provided it is taken as an accumulation period and not as a period of investment." I confess I do not understand the meaning of this distinction, since in the context output value and input value represent permanent time streams, while input is regarded as "provision for maintenance or as payments for the other agencies co-operating with the particular capital good . . . or as including elements of both" (*ibid.*, p. 56). The "period" clearly cannot refer merely to the time during which the capital stock is accumulated (which is the sense in which the term "accumulation period" is generally used).

[20] If a is the value of the annual input, b the value of the annual output, t the average period sought, their relation will be given by the equation

$$a(1+i)^t = b. \tag{1}$$

The rate of interest in question, however, is the investment's "real rate of return," which is given by the equation

$$i = \frac{b-a}{C}. \tag{2}$$

where C is the value of current services needed to produce (or reproduce) an "investment" capable of yielding an output stream b at an input stream a. Since the production of resources also takes a certain time, this construction cost will itself include an element of interest. This, however, causes no logical difficulty; for the construction cost (including interest) will still have a unique value if we impose the further condition that interest during construction must be identical with the interest earned on the investment itself. In other words, given the input of all dates (including the series of initial inputs, representing

of the investment is merely the difference between the two and since, under our assumptions (i.e., constant returns to scale), every unit of capital in that investment is assumed to earn interest at the same rate, the relation between output value and input value will also determine the relation between "construction cost" and "annual maintenance cost." For investments which are continuously maintained at an even rate of time, the degree of roundaboutness can be measured by the ratio of the initial or construction cost to the annual maintenance cost (assuming that the expected future prices of productive services are the same as present prices).[21,22] The "law of roundaboutness" then

"construction") and the outputs of all dates, the rate of return on the investment will be uniquely determined.

[21] We use the expression "the degree of roundaboutness" rather than "the investment period," since the ratio of construction cost to annual maintenance cost gives us an *index* to the period of investment, rather than the period itself. It will correspond to the average period (as defined above) only if the rate of interest is small and compound interest can therefore be neglected. Neglecting compound interest, the above equation (1) becomes

$$a + (a[1 + i] - a)t = b,$$

from which

$$t = \frac{b - a}{ai}.$$

But

$$C = \frac{b - a}{i} \text{ (from (2))};$$

$$\therefore \quad \frac{C}{a} = \frac{b - a}{ai} = t.$$

This is also the definition of the "time spread of the investment" given by Mr. Boulding ("Time and Investment," *op. cit.*, pp. 212–213) who appears to have reached it by a different route; and also of Mr. Smithies ("The 'Austrian' Theory of Capital in Relation to Partial Equilibrium Theory," *op. cit.*, p. 81). Its merit is that it enables us to make use of the concept also in those cases where the lifetime of individual capital goods comprising the investment cannot be evaluated.

In the case of our three types of house investments, the relation of annual output to annual input will be 225/100, 273/100, and 312.5/100 respectively, while the ratio of "initial cost" to "annual maintenance" cost will be (approximately) 2000/100, 2800/100, and 3600/100. Since on account of compound interest the value of individual houses does not diminish at an even rate in time, the "replacement cost" of a stock of houses with an average age equal to half their lifetime will not equal half their total cost of construction, but will be higher than this amount. This ratio will therefore only approximate to Böhm-Bawerk's "average period of production" (= half the lifetime of the houses) when the rate of return is so small that compound interest becomes negligible.

[22] The "annual maintenance cost" of a resource (or good) includes the value of *all* services consumed in producing whatever is regarded as the output stream of that particular resource. It is determinate therefore only if the output

simply says that it is always possible to reduce annual maintenance cost by increasing initial construction cost, in producing a given permanent output stream.

Now, according to Professor Knight, this concept of the investment period, or "degree of roundaboutness," is without significance for capital theory; for "the average investment period and the quantity of capital may perfectly well be affected in opposite ways."[23] The argument, if I rightly understand it, could be summarised as follows: The optimum degree of roundaboutness, on any single investment, is the one which maximises the rate of return on that investment. A change in the quantity of capital could only lead to a shift in the optimum degree of roundaboutness by affecting the relative rates of return on different degrees of durability. It is usually assumed that this will be the case because an increase in the supply of capital will lead to a fall in the rate of interest. But in the case of "continuous maintenance" the *rate of return*, on any single investment, will be independent of the rate of interest. It is only by assuming that the amortization quotas are accumulated at some "outside" rate of interest that this "internal rate" will be affected; in which case a given fall in the rate of interest would reduce the return from less "durable" investments to a greater extent. In the numerical example we have given above, the reduction in the interest rate to 4 per cent would make Type II houses more profitable than either of the other two types. But this method of calculation is obviously mistaken since it overlooks the fact that, by reinvesting the amortization quotas in the same uses, a much higher net return is obtained than by reinvesting them at the current interest

stream of the particular good is regarded as given. Since, however, the resources themselves can only be unequivocally defined by their output streams, this problem ought to cause no difficulty. To elucidate our concept by an example: if the output stream of certain boot-manufacturing machines is regarded as a certain quantity of machine services per unit of time (assuming that these services are capable of physical measurement, in terms of machine-service-hours, like labour-hours), the "annual maintenance cost" or "input value" of those machines will consist of the expenditures—in the form of upkeep and replacement—continuously incurred in securing a permanent flow of these services. If, on the other hand, not "quantity of machine-service-hours, per unit of time" but "quantity of boots per unit of time" is regarded as the output stream of those particular machines, "the annual maintenance cost" will include, in addition to the above, also the cost of the services of the factors (labour, etc.) normally regarded as co-operating with the machines in producing the boots. The ratio of construction to maintenance cost—which, perhaps, should more properly be called *the ratio of the initial input to the annual input flow*, the former, as distinct from the latter, being a singular expenditure which is incurred only once, at the beginning of the investment—will of course be different in the two cases: but so will the "investment period," if measured in any other manner.

[23] *Ibid.*, p. 45.

rate outside. It is not true, therefore, that a fall in the interest rate would make it profitable to shift to more durable houses. In the above example, the least durable house (Type I) has the highest real rate of return—6.35 per cent—and so long as the price of room service and the rate of wages remain the same, this is the type that will be preferred, irrespective of how much the rate of interest might fall.

An increase in the quantity of capital, therefore, will not change the "degree of roundaboutness" involved on *already existing* investments; and there is no reason to suppose that this "degree of roundaboutness" will be higher on new investments than the average on already existing capital goods. What happens when the rate of interest falls is that investments whose real rate of return was lower than the previous interest rate become profitable. More houses will be built. But the houses which have only just become profitable on account of the lower rate of interest need not be "more durable houses"; they may be houses with a different quality of room service. It is the relation between net return and cost of construction which must be lower. But the kind of houses which have a lower net return may very well have a lower ratio of construction cost to maintenance cost and thus a lower "period of production." The two are not related to each other at all—durability, as Knight contends, is merely one of an "infinite number" of considerations that affect the net return of investments.

III

Before we proceed to a criticism of this argument, we might attempt to piece together these various aspects and give a general picture of the world as Professor Knight sees it. It consists of a collection of resources, which, like heavenly bodies, emanate light and absorb light. All these resources have to be "maintained"; i.e., they all absorb a quantity of services at every unit period which is the absolute condition of their continuing to radiate another stream of services, which is their "output." No distinction can be made between maintenance and replacement, or even between production for immediate consumption and production for "maintenance"—or future consumption—since all that we know is that during a certain period a certain quantity of all kinds of services have been "put in" (into each particular "resource" or "factor") and a certain other quantity of services has been "put out." It is impossible to say "how much" of the input served to produce the immediate output, and how much served to maintain the resource itself. And since, in a well-organised competitive world, for each particular resource both input stream and output stream must be constant, per unit of time (if the ruling prices are

expected to remain in operation),[24] the question itself is meaningless. Looked at in one way, all production is "instantaneous"—if the input stream is regarded as "producing" the output stream. If the resources themselves are regarded as producing the output stream, all input is to be regarded as producing output in an indefinite future. The output stream of all resources in so far as they do not directly consist of consumption service and in so far as they are not actually creating some *additional* resource—must therefore be input or "maintenance cost" for some other resources. Even consumption can be looked upon as the input of the resources called "labourers." Not all consumption, of course. For on the one hand labourers' consumption falls short of total consumption by the consumption of the owners of other resources —on the other hand, the labourers' consumption must itself include the net return from the investment of owning themselves. This difference (property owners' consumption plus the difference between labourers' income and maintenance cost) can be regarded as the "net return" from the whole system. It is precisely the extent to which all inputs fail to cancel out all outputs.

In a growing system some of the service stream (of all types of resources) will also be engaged in producing further resources. To the extent that such services are obtained by reducing the input-stream of other resources—and this is the only way of obtaining them if a world of "full employment" is contemplated—these other resources, will, for the period of construction of the new resources, be "undermaintained"—their input stream will be temporarily reduced. Not all the resources "lent" will be repatriated, of course, at the end of the construction period. Some of them will permanently remain with the new resources, as their permanent input flow. This deficiency, however, will be more than offset by the output stream from the new resources, which directly or indirectly will also help to maintain the old ones.[25]

As the quantity of capital is increasing, the rate of return falls, since this implies the adoption of progressively inferior investment opportunities. It is at the margin of investment that the rate of interest

[24] I believe this assumption underlies the whole of Knight's analysis. When he mentions "perfect foresight" he uses this word in a different sense from the one in which Professor Hayek uses the term. Professor Knight, I believe, merely implies that the markets are sufficiently perfect to adjust themselves immediately to any given change—they are "Walrasian" markets. It is "perfect foresight" only under the static assumption that no further changes occur in the future.

[25] The whole situation is analogous to the case of a hydroelectric plant, which lends part of its water power for the construction of another plant. Once the new plant is constructed, the old plant's power will no longer be required except for "maintenance," which is a small proportion of the construction cost and, if I rightly interpret Professor Knight, could easily be less than the additional net output of the new plant.

is determined; capital quantity itself is a "marginal concept." Accumulation implies the conversion of current income into additional streams of permanent income; it implies an increase in "resources" in general, in the capacity to produce output streams, and in this sense *every* addition to the stock of capital should be considered as a permanent improvement. Accumulation requires abstinence (in the sense that abstaining from a part of the current product is the price of creating an additional output stream) but there is no "waiting period" involved in the maintenance of a given stock of resources,[26] and, since the services of all resources equally contribute to the creation and maintenance of each other, no definite meaning can be attached to the term of an investment period itself. This concept is in any case *irrelevant;* for even under the most favourable assumptions it could not be substantiated that an increase in capital will necessarily imply the adoption of "lengthier" processes.

I am not sure whether this brief picture does justice to Professor Knight's views. But if it is a correct interpretation of his theory, it fails to account for a number of factors which it is the fundamental task of a theory of capital to explain. In the first place, it does not explain how the rate of return, on different investments, is kept at a level of equality. Under the conditions postulated, the rate of return should correspond in equilibrium to the current rate of interest not only on the marginal unit of investment, but on all units. It can be argued that "inframarginal" investments will earn rents which, in terms of money costs, will equalise this difference; but then the question still arises: why should "rents," if they arose, not be eliminated by competition? In the second place (and this is closely linked up with the first) it

[26] Among Austrian theorists, the "waiting period" is sometimes measured by the extent to which current consumption has to be reduced (below some technical maximum) in order to permit the maintenance of the existing stock of capital, i.e., in order to secure the continuance of the same rate of consumption permanently. Now it is perfectly true that at any time, given the technical composition of the system, the rate of consumption could be stepped up a certain extent if all productive services were devoted to producing for immediate consumption—given the length of time for which the increased rate of consumption-output is supposed to last. But the extent to which this can be done will depend on the type of capital goods used as well as on their quantity; and it is quite possible that with an increase in capital, the possibility of expanding consumption by not maintaining capital goods should decrease rather than increase. In any case, the extent to which this can be done will certainly have no relation to the *value* of capital in terms of current income, except in those simple cases where the capital consists exclusively of circulating capital, physically homogeneous with the final product. (E.g., if capital consisted of the stock of grain annually reinvested—in the form of seed and advances to labour—the quantity of consumption could be expanded in precisely the same ratio as the value of the capital stock in terms of the annual product.)

does not really explain why an increase in capital should lead to a fall in interest. To say that resort must be had to inferior investment opportunities does not in itself meet the problem. Diminishing returns necessarily presuppose the existence of some "fixed factor" as their cause; and there is no room for such "fixed factors" if we regard, as Professor Knight apparently regards, capital accumulation as an increase in the quantity of resources in general. In the last place, this theory contributes little to an explanation as to how *interest as a distributive share* is determined, along with other distributive shares. The great merit of the Austrian capital theory—at any rate of Wicksell's version of this theory—is that it explains the interrelation between wages and interest; and thus makes it possible to extend the general marginal productivity theory so as to include capital. So far as this problem is concerned, the critics of the traditional theory can hardly be said to have offered an alternative explanation.

We shall attempt to demonstrate in the following that the crucial argument concerning the irrelevance of the "law of roundaboutness" ignores the all-important effect of a change in the quantity of capital on price relationships; and that an interpretation can be given to the theory which allows it to survive most of the other criticisms that have been brought forward. Finally we shall endeavour to show that the "law of roundaboutness" itself is merely a derivation from the general law of nonproportional returns; while the Austrian view of capital merely implies an attempt to measure the quantity of variable resources by the average productivity of the services of "fixed" resources, which is possible so long as the latter are homogeneous in kind and the composition of the final output stream can be considered as given.

IV

1. In the first place, let us go back for a moment to the question of the definition of resources. Here Professor Knight appears to have overlooked one distinction which survives the strictures levelled against the traditional classification. Even if all resources require to be maintained and the services of all resources contribute to the production of new resources, it is still not true that *all* kinds of resources can be *produced*. It is not possible to produce "land"; and, in a capitalist economy which no longer knows the institution of slavery, it is not even possible to "produce" labour. The quantity of labour, through a change in the birth rate, can certainly be increased, but to regard this quantity as being a function of saving or the rate of interest is turning an analogy into a falsehood.

If the services of producible resources provided "perfect substi-

tutes"[27] for the services of the nonproducible resources this difference would not constitute a "relevant economic fact"—the prices of the services of nonproducible resources would be entirely governed by the services of produced resources. In reality, however, the services of capital goods provide merely an imperfect substitute to services of labour; the one can be substituted against the other in any sort of production only at continuously increasing marginal rates of substitution. Thus even if the distinction between "permanent" and "nonpermanent" resources or between "original" and "produced" resources were untenable or irrelevant, there is still a distinction to be drawn between "producible" and "nonproducible" (or rather, "augmentable" and "nonaugmentable" resources).

Given this distinction, we must immediately make note of another factor, which in this paper has so far been left in the background: that in a position of equilibrium, assuming perfect competition, the value of producible resources must always correspond to their cost of reproduction (to the value of the quantity of services needed to produce another, "identical" resource). The value of nonproducible resources, on the other hand, need not conform to any such criterion simply because they have no costs of reproduction.

Now, what Professor Knight's own theory has not explained—or at any rate the present interpretation of his theory has not explained so far—is the problem, how this correspondence between the value of producible resources and their costs of reproduction is achieved, or if achieved, how this correspondence is again re-established, once equilibrium has been for any reason disturbed. A fall in the rate of interest, e.g., will raise the discounted value of *all* future income streams, and thus the present value of all resources whose ownership can be bought and sold (that is to say, all resources except labour). Moreover, if it is assumed that all resources are "continuously" maintained, it must raise the market value of all investments in the same proportion. If their value was previously equal to their costs of reproduction, they will now exceed these costs by the proportion which the fall in the rate of interest bears to the new rate of interest. How will this correspondence be re-established?

2. In order to analyse the interrelation of different factors let us return to the simplest hypothetical situation, where the stock of capital consists of houses which are built exclusively by labour, while "room-years" represent the only kind of consumption good. In order to avoid monetary complications which are not relevant in the present discus-

[27] In the sense of their having infinite "elasticities of substitution" with the services of the other resources, i.e., that this rate of substitution did not vary with the proportions in which they were combined.

sion, we might also assume that "room-years" serve as a *numeraire* in terms of which debts are contracted, wages are paid, and property is valued. In this society "savings" imply a desire to convert current income ("room-years") into "houses"—in other words, an increased desire for "holding" houses. If this increased demand can be satisfied by an increased supply (when, e.g., unemployed labour is available for additional house building) there need be no change in the value of houses in consequence. But if *all* the labour is already engaged in building (or rather "replacing") houses, it is the value of houses that will rise (which is merely another way of saying that the rate of interest, in terms of room-years, will fall); and, as the value of houses rises, wages will rise. For the value of existing houses cannot be higher than their costs of reproduction; and a rise in costs of reproduction must imply a rise in wages.

Alternatively one might say that saving first leads to a fall in the room-year rate of interest (which is "determined" in the annuity market), this creates the rise in the value of houses, which in turn increases wages. The rise in wages increases construction costs; but it will also reduce the value of houses (i.e., below their new level, which they reached after the fall in interest). For the rise in wages, by raising expected future wages, increases maintenance costs, relatively to gross incomes (input values relatively to output values) and thus reduces the "net incomes" on the basis of which capital values are calculated. Thus, while costs of construction rise, capital values fall, and "somewhere in the middle" they again meet, thus bringing the movement to an end. In either case, it is the change in wages which brings the real rate of return on individual investments into equality with the rate of interest.

It would seem to follow from this that in this society "savings" merely resulted in a transference of income from the capitalists to the labourers.[28] There would be no increase in aggregate real income; and (save for changes in relative demand arising out of changes in distribution) there would be no changes in composition. In particular, it is difficult to see how investment opportunities which were previously ultramarginal (which were previously not adopted because their real rate of return was lower than the prevailing interest rate) would, as

[28] This transference would not be temporary, but permanent, (even if "savings" were temporary). For it would be financed, so to speak, out of two sources: first, the increase in the supply of capital, coming from the savers; second, the reduction in interest (in the return on investments) which the increase in the supply of capital creates (and which would thus be shared equally by all capitalists). The reduction in the interest rate, following upon a given increase in capital, would be precisely such as would enable the same transference of real income per time unit permanently as the volume of savings (per time unit) which was originally responsible for it.

a result of savings, become inframarginal. For the rise in wages would have offset the effect of the reduction of interest; and in the new situation, they would still be below the margin of profitability. Continued capital accumulation under such circumstances would merely lead to the complete expropriation of the capitalists, by reducing the rate of interest to zero and making the value of annual labour input identical with the value of room-year output.[29]

3. But fortunately for the capitalists this will not be so—not even under our rigid assumptions. For the rise in wages in terms of house-room creates something which by itself tends to check the tendency of the level of wages to rise and the income from capital to fall. It necessarily increases the optimum degree of roundaboutness. room-years ?

Let us return to our numerical example of the three types of houses and see how their respective rates of return will be affected by varying increases in wages. Since the rise in wages must always be such as to equalise the rate of interest with the real rate of return, this will also show the level of wages corresponding to different rates of interest (represented by the italicised figures):—

Increase in Wages (per cent)	Real Rate of Return (per cent) of		
	Type I	Type II	Type III
0	*6.35*	6.2	5.9
10	5.42	*5.45*	5.24
20	4.65	*4.83*	4.68
50	2.69	3.29	*3.37*

We can see from this that not only does Type II become the most profitable investment if the increase in wages is 10 per cent, but the differences in profitability, expressed as a percentage, continuously increase with every increase in wages.[30] Assuming that there is a con-

[29] This sounds rather like a rehabilitation of the classical theory of the Wages Fund—which in a sense it is meant to be. If conditions were postulated under which an increase in the supply of capital would *not* lead to an increase in aggregate real income (when, e.g., the technical coefficients between "capital" and labour—the services of produced and nonproduced resources—are fixed and the quantity of labour is given) the supply of capital would determine—in a linear fashion—the rate of wages. There is no reason to assume that in such a society the rate of interest will be necessarily zero—it will be determined at the point where the demand for "annuities" (in exchange for current income) is equal to its supply. (The rate will be zero only if at any positive rate the demand for annuities exceeded the supply.) The rate of interest thus determined will determine the level of wages and the share of labour in the product.

[30] In the above example, the changes appear numerically slight (relatively to the changes in wage rates) but this is only because the maintenance costs, in the examples shown, were already very low in relation to the construction cost. Generally speaking, the numerical change in relative profitability for a given

tinuous range of alternatives and not merely three distinct types of durability there must be a shift in the optimum ratio of construction cost to maintenance cost (or input volume to output volume) as soon as the price of input units rises relatively to the price of output units. This shift can be thought of as being brought about (for the "representative enterprise") either by a reduction of present "output" with a view of increasing the future rate of output (the input stream remaining the same) or a reduction in present output with a view of reducing future rate of input; or, finally—since the input flow is subject to diminishing returns in terms of output flow—simply a reduction in the permanent rate of input which is followed by a less-than-proportionate reduction of the permanent rate of output. In all of these cases there will be a reduction in the permanent input flow per unit of output flow; which in turn will have three different consequences: In the first place, it damps down the fall in the value of investments, brought about by a rise in wages; since the increase in maintenance cost will no longer be proportionate to the increase in wages. In the second place, it increases the "costs of reproduction" of house investments more than in proportion to the increase in wages (since maintenance costs can only be reduced by increasing construction cost) and thus closes more rapidly the "gap" between the value of investments and the costs of reproduction, caused by a given increase in the supply of capital. (In other words, it closes the gap with a smaller increase in wage rates than otherwise.) All this can also be expressed by saying that the existence of Type II houses as alternative to Type I houses prevents both the rate of interest from falling, and the level of wages from rising, so much—following upon a given percentage increase in "free capital"—as they would have fallen, or risen, had Type II houses not been available as an alternative. In the third place, it creates an increase in the permanent supply of house room, which otherwise could not have taken place, as a result of a fall in the interest rate.[31] If in the above example we further assumed that there was only a single kind of house room in existence (that given in the example)

increase in wages will be greater, the higher is the ratio of maintenance cost to construction cost—(the influence on relative profitability of changes in the interest rate in the case of "discontinuous maintenance" will be *per contra* the more noticeable the lower is this ratio)—and greater the higher is the real rate of interest. With continued increases in wages, the differences generally increase in a diminishing proportion.

[31] Furthermore if we assume that the "degree of roundaboutness" for different types of room-years is different, the rise in wages will change their relative rates of return. For a given rise in wages will affect the rate of return all the more the higher is the ratio of annual maintenance cost to construction cost. The re-establishment of equilibrium (i.e., equalisation of the rates of return) will then require, in addition, a relative fall in the prices of the services of more "durable" resources and a relative expansion of their supply.

the change over to Type II investments from Type I investments will ultimately have increased the volume of available room-years in the ratio of 100(273–225)/225, i.e., by 20.88 per cent. This, divided by the quantitative increase in the investment period, which is involved in this change-over, should give the "marginal productivity of waiting" according to the Jevons' formula, to which the rate of interest must correspond at the point where the two types of investments are equally profitable.[32]

Thus, given the available quantity of labour and the productivity function of capital (the extent to which maintenance cost per unit of output can be reduced by a minute increase in the ratio of construction cost over maintenance cost), the rate of interest determines the relative price of labour service and consumption service. This price ratio in turn determines the "average investment period," i.e., the degree of roundaboutness which maximises the yield of investments. Alternatively, the increase in the supply of capital determines the extent to which the degree of roundaboutness will be changed by changing the ratio of the price of input units relatively to output units, which in turn determines the rate of interest, since in equilibrium the rate of interest must be equal to the "real rate of return" on investments.

All this is merely a simplified and somewhat loose account of the Wicksellian version of the Austrian theory, first put forward in the *Über Wert, Kapital und Rente,* and later in the *Lectures on Political Economy,*[33] and adapted to the case where all capital is "permanently and continuously" maintained. It differs from the Böhm-Bawerkian theory chiefly through the analysis that for the individual entrepeneur the optimal investment period is determined by the production function and the existing price relationships (which are given to him); while the supply of capital "determines" the investment period by determining the ratio of output prices to input prices (i.e., of a unit of consumption service to a unit of labour service).[34]

[32] The two types of investments become (approximately) equally profitable at a wage increase of 6 per cent at which both yield 5.8 per cent. At this rate the "compound investment period" (calculated according to the formula in footnote 20 above) will be 14.73 years for Type I and 17.85 years for Type II. The net increase will therefore be 3.42 years and the "marginal productivity of waiting" $20.8/3.42 = 6.2$ per cent, i.e., approximately the same as the rate of interest. (An exact equality could only result if very small changes were contemplated.) Since, however, in these cases, the "investment period" (in terms of years) can only be evaluated if the rate of interest appropriate to the situation is already known, the concept of the "marginal productivity of waiting" does not seem to be particularly helpful.

[33] Cf. *Lectures,* Vol. I, Part II, Sect. D, "An Alternative Treatment of the Problems of Interest and Distribution," also Appendix 21, "Real Capital and Interest." Cf. also Wicksell's *Finanztheoretische Untersuchungen,* pp. 22–41.

[34] Cf., e.g., *Finanztheoretische Untersuchungen,* p. 33 (my translation): "Given

V

So far we have merely attempted to vindicate the traditional capital theory under the simple assumption that the capital of the world consists of houses produced exclusively by labour; that there is perfect competition, static foresight, and the absence of uncertainty. The real world—for the purpose of the present discussion—differs from this, apart from the last three assumptions, in three important respects: (i) that the maintenance of capital does not have the character of "replacement" of units at definite intervals but rather that of continuous repairs; (ii) that the services of labour are not at all invested in capital but partly co-operate with the services of capital goods in producing consumption services, i.e., the labour force itself is divided, to use Wicksell's expression, between "free" and "invested" labour; (iii) that capital goods are not produced exclusively by the services of labour but also by the services of other capital goods, i.e., the services of capital goods themselves help to produce (or "maintain") each other. How far do these facts modify our results?

(i) The first of these points can be treated briefly. Whether "maintenance expenditure" consists of definite replacement of physical units or merely of repairs, the ratio of initial cost to annual maintenance cost will still provide a measure of the "degree of roundaboutness"; and so long as it is still possible to reduce the annual maintenance charge, of a given service stream, by increasing the initial construction cost, it will still be true that the price ratio between output units and input units will determine the optimum relation between construction cost and maintenance cost, which, in turn, will determine the rate of interest. It will not be possible, of course, to associate a definite "investment period" with the input of any *particular* period; but this, as we have seen, is hardly legitimate in any case, unless the whole contribution of the input of a particular period accrues at some given date in the future (as, e.g., with the storage of wine), which is only true in certain specific cases.

(ii) The second point is more serious. It affects our previous analysis in two ways: (a) In the first place, it is clear that if a part of the labour supply is co-operating with existing equipment in producing current output, simultaneously with savings a certain quantity of labour will be "released" for employment in new construction. If instead of houses we had taken the less unreal example of machines co-operating with

---

the general postulate of Böhm-Bawerk's theory [i.e., the law of roundaboutness] one would think at first that the capitalist always aimed at a steadily longer investment period of his capital—at any rate once the loss of interest during the transition period can be neglected. This, however, will by no means be the case; for any given level of wages, there is always an optimal length of the investment period."

labour in producing bread, it would have been at once obvious that savings will not merely increase the demand for "holding" machines, but will also reduce the demand for bread. Corresponding to the increase in the demand for labour in machine-making, there will be a released demand for labour in the making of bread.[35] If machines are produced exclusively by labour, while "bread" is produced partly by labour and partly by machines, there would still be an increase in the *aggregate* demand for labour. But if "labour" and "machines" co-operate in the same way in producing new resources as in producing final output—and this is what Professor Knight's first point really amounts to—there need be no net increase, as far as the *creation* of new capital goods is concerned, either in the demand for labour services or in the demand for machine services. There could thus be an increase in the number of machines even without a rise in wages. It would be wrong to conclude, however, that this would invalidate our previous conclusions. For once the new machines are in existence and "saving" correspondingly ceased, they will require some additional labour for their maintenance and operation which they can only get by reducing the quantity of labour employed in combination with the previously existing resources. This in turn (if machine services are merely an imperfect substitute for labour services) will increase the price of labour services, relatively to other services (which is merely another way of explaining that the relative increase in "other services" increased the relative scarcity of labour services), it will reduce the quantity of labour input per unit of bread output (by reducing either the labour embodied in, or the labour co-operating with, a unit of machine service, or both), which in turn implies an extension of the degree of roundaboutness and a fall in the rate of interest. It still remains true that it is the rise in wages, in terms of final output, which causes the fall in the rate of return; a fall which would be more severe if it were not possible to offset partly the effect of the rise in wages by extending the degree of roundaboutness.

(b) This brings us to the next point in this connection: the question of durability. We have already mentioned earlier[36] that the input

[35] The reason why this has been apparently overlooked (by the classics and in Wicksell's treatment, cf. esp. *Lectures, op. cit.*, pp. 148–149) is due to the assumption that what is saved is the product of past labour and not of current labour, so that the current demand for labour is independent of current consumption; depending only on the current supply of capital. (This is the meaning, e.g., of Mill's statement that the "demand for commodities is *not* a demand for labour"). This again is true if (a) the unit of account is fixed in terms of the final product, so that changes in current consumption do not affect the profitability of investment *via* price expectations; (b) all labour is "invested labour"—as, e.g., in the case of an agricultural community, whose labour requirements consist mainly in sowing seed for the following harvest.

[36] Cf. footnote 22 above.

stream (and thus the ratio of initial input to annual input) of resources will depend on how one defines the "output stream" of resources. In the example just given, either the "quantity of machine services per unit of time," or the "quantity of bread per unit of time" can be regarded as the output stream of the machines. In the first case the "input stream" will consist only of expenditures incurred in the upkeep and replacement of the machines (Wicksell's "invested labour"). In the second case it will include, in addition to the above, also the labour normally regarded as co-operating with the machines in producing the bread. According as the first view is taken or the second, we shall have two different measures of the "degree of roundaboutness." Only the first of these can be regarded as an index of the *durability* of capital goods. But only the second will be necessarily correlated with the quantity of capital.

It is only in so far as the proportion of invested labour to co-operating labour remains constant when the aggregate quantity of capital changes that the degree of roundaboutness will necessarily increase in both senses. And although this follows from Wicksell's analysis of the problem[37] there seems to be no reason that it should be the case if the possibility of a change in the *character* of the machines is taken into account.[38] An increase in the quantity of capital available might even lead to the introduction of *less durable* rather than more durable equipment, if only this equipment is more "automatic" (in the sense

[37] Cf. *Lectures*, Appendix 2, pp. 278 ff., esp. 287–288.

[38] Wicksell's argument could be summarised as follows: Let us suppose that in the beginning the increase in capital only leads to an increase in the number of machines, of the same type as those already in use. This will imply that the amount of invested labour increases and the amount of "free" labour is reduced; which in turn will necessarily raise wages and reduce the price of the services of machines. The rise in wages, as we have seen before, makes it profitable to extend the lifetime of machines, which in turn will imply a reversal of this process: the amount of free labour will increase and the amount of invested labour will be reduced. On Wicksell's assumption this must continue until both regain their former proportion. Meanwhile "the labourers lose part of, but not all of, their recent increases in wages and the capital goods regain part of, but not all, the value they have just lost." (*Ibid.*, p. 288.)

It is quite possible, however, that as a result of the rise in wages, it becomes profitable to introduce not more durable but more automatic—and even less durable!—machines and in consequence there will be a further increase, rather than a reduction, in the amount of invested labour. It is often thought that machines which are both *more* efficient and *less* durable will be preferred irrespective of the quantity of capital. That this is not the case, can best be elucidated by a simple example. Let us assume, e.g., that bread can be manufactured by two different processes. The first involves machines which require an initial expenditure of 1000 units of labour and an annual maintenance of expenditure of 10 units (per unit of bread, per year). These machines will need in addition 50 units of labour to operate them. The second involves machines which require

of requiring less labour to operate it) than the previous equipment. It is not true therefore (except in the special case, like houses, where all the labour used is invested labour) that the increase in the quantity of capital will *necessarily* lead to an increase in "average durability," or that it will lead to the making of "goods of still greater durability in place of those produced before."[39] It *could* imply the opposite of these things. It must necessarily increase the "degree of roundaboutness" involved in producing final output (if co-operating labour and invested labour are taken together); but this is *not* the same thing (except in the special case where the amount of co-operating labour is zero) as an increase in the average durability of capital goods.

(iii) The last point—although it is the one most frequently emphasised by other critics[40]—does not, in our view, affect the theory any more than it has already been affected by previous considerations. It is perfectly true that at no stage of the production process is labour exclusively employed—the services of different types of resources contribute to the "maintenance" (or production) of each other; the output stream of resource A might be the input stream of some other resource B, whose output stream in turn forms part of the input of A. But this does not imply that this "circularity" in production is complete: this would only be the case if consumption itself could be regarded as part of the system's "input."[41] Now all "outputs" (of resources other than labour) which are *not* consumption services must be simultaneously inputs in some other resource. Similarly, all inputs, in so far as they do not consist of labour service, must be the outputs of some other resource. Therefore all outputs which are not consumption service and all inputs which are not labour service, exactly cancel

---

an initial expenditure of 1500 units and an annual maintenance expenditure of 40 units; but these machines being much more "automatic" only require 10 units of labour to operate them (all per unit of bread, per year). The ratio of initial cost to annual maintenance cost in the *first* sense will be 1000/10, 1500/40 respectively, in the two cases. In the second sense, it will be 1000/60, and 1500/50 respectively. Now, if the price of labour in terms of bread is unity, obviously the first of these methods is preferable to the second—since it will yield a return of 4 per cent while the second yields only 3 per cent. If, however, the price of labour rises, say by 50 per cent, the second method will become preferable to the first; since in that case, the yield on the first method will be reduced to 0.5 per cent while the yield on the second only to 0.83 per cent.

[39] Machlup, "Professor Knight and the 'Period of Production'," *op. cit.*, p. 590; and Hayek, "The Mythology of Capital," *op. cit.*, p. 213.

[40] Cf. Joseph and Bode, "Bemerkungen zur Zinstheorie," *op. cit.;* Nurkse, "The Schematic Representation of the Structure of Production," *op. cit.*

[41] It is possible, of course, to regard that part of the labourers' consumption which is necessary to maintain this productive capacity intact, as the "input" of labour as a factor of production. But only in a slave state would this magnitude have an economic significance.

each other out, if the input streams and output streams of individual resources are added together.[42] By defining the "net output" of resources as the volume of consumption we thereby also necessarily define their "net input" as the quantity of labour.[43] So long as the quantity of annual labour service remains constant with variations in the quantity of capital, and so long as the quantity of no other type of services remains constant, there will be a unique correlation between the rate of interest and the amount of labour input per unit of final output—or, if you like the rate of interest and the average investment period of the services of labour. For, as I hope to show in the next section, the "investment period" of a factor necessarily varies with its average productivity, once it is assumed that the factors themselves have a cost of production and not only the final products.

VI

For a proper understanding of the nature of capital and interest one ought to start by analysing the conditions of equilibrium in a society where *all* goods are capital goods, i.e., where "original" or non-augmentable resources do not exist at all. It is rather unfortunate that, following Böhm-Bawerk and his school, we have been generally accustomed to start with a more specialised set-up, with the picture of Robinson Crusoe engaged in net-making. This Crusoe-approach makes it unnecessarily difficult to single out features which are merely the property of a special case from the demonstration of general principles. Had the analysis started with the "general case"—by imagining a society where *all* resources are produced and the services of all resources co-operate in producing further resources—a great deal of the controversies concerning the theory of capital might not have arisen. As we shall see, it will be much easier to get back from this world to Böhm-Bawerk's world than to make the journey in the opposite direction.

Let us imagine, then, a society where "machines" and "slaves" are the only scarce resources, whose services are required equally for the production of each other and for the production of bread.[44] The owners

[42] Cf. also the "analysis of interactions" in Fisher's *Theory of Interest*, pp. 18–22.

[43] This really follows from selecting "labour" as being distinct from other resources, in which case the input of *all* resources other than labour will consist of labour service. It would also be possible to regard some other factor—"land" —in the same way: in which case the input of all resources (including labour under this head) would consist exclusively of land service. The reason for regarding "labour" as distinct, is twofold: (a) that it is the ownership of labour which is nonalienable and in consequence has no capital value; (b) that it is the quantity of labour service which can be regarded as a constant with respect to "saving." Cf. also the next section, below.

[44] I.e., there is a production function for machines, whose variables are machine service and slave-labour service, a similar one for slaves, and yet another for bread. If we strictly adhered to the terms of our example, it should be added

of slaves and machines (the entrepreneurs) will, under these assumptions, have essentially three degrees of freedom: (1) they can vary the proportions in which the services of machines and slaves are combined in the production of bread; (2) they can vary the proportions in which the machines and slaves themselves are produced, or reproduced; (3) they can decide how much of the "net output" of any period (i.e., the quantity of bread production compatible with maintaining the stock of slaves and machines intact) should be set aside to increase the permanent stream of bread output in the future.

Assuming perfect competition and constant returns to scale, the entrepreneurs will (individually) combine the two factors in such proportions as to maximise the output of a given outlay; and they will tend to produce the factors themselves in such proportions as would maximise the rate of return on a given investment (all in terms of "bread"). Assuming that the law of diminishing productivity operates throughout (i.e., that there is an increasing marginal rate of substitution between machine services and slave-labour services, in the production of bread, machines, and slaves) the problem will have a unique solution. Given the cost function of machines, slaves, and bread, there will be only one proportion between machines and slaves which will maximise the yield of capital: the proportion at which the value of both machines and slaves (calculated by discounting at the same rate their expected net income) is equal to their respective costs of reproduction.[45] It is this yield which in turn will determine the rate of interest. (All this can also be expressed by saying that the yield on capital will be maximised when the real rates of return, on machine investments and slave investments, are equalised.) This rate will represent at the same time the system's "maximum rate of growth": the rate at which the stock of resources would increase, per unit of time, if consumption is reduced to zero and the services of all productive resources were devoted exclusively to their own production.

---

that the services of machines and of slave labour are directly required only for the production of machines and of bread. Only "bread" is required for the production of slaves. But bread in turn represents a certain quantity of machine services plus labour services, combined; so that we can say that the services of both resources are needed for the production of both resources.

[45] If there is a relative increase in the number of machines, and a consequent fall in the yield of machine investments, this would not imply an equivalent fall in the yield of "capital"—as it does in our own society—since the fall in the yield of machines would be largely offset by the corresponding increase in the yield of slaves. But on account of the law of diminishing returns it could never be so offset *entirely* (and vice versa if there is a relative increase in the number of slaves). Thus there will be only one ratio of investment in the two factors which equalises the real rates of return on these two types of investment and this will necessarily be also the arrangement which maximises the return per unit of bread.

Thus both factors will yield a "net product"—i.e., the specific productivity of their services will be greater than the costs of production of these services—and the rate of return merely denotes the size of this excess, per unit of time, as a percentage of the cost. Since this "real productivity," and thus the real rate of return, on any resource will depend upon the relative scarcity of the services of that resource, and since the proportions of the factors are variable, investment will tend to get distributed in such proportions as would equalize the rate of return on all lines of investment.[46] Once this proportion is achieved, capital accumulation or decumulation (in the absence of a change in technical knowledge) will leave the rate of interest unaffected. How rapidly capital will be accumulated will depend, of course, on the rate at which people are willing to save at the given rate of interest; but no amount of capital accumulation could change this rate.[47]

In this society there will be two distinct "investment periods" which cannot be combined for the purposes of an average, since they are alternative ways of describing a single situation. We might either represent the entire bread output as the product of machines whose input consists of slave-labour service; or we might represent the entire bread output as the product of slaves whose input consists of machine service. The average investment period of the services of slave labour will depend on the ratio of the value of the entire labour input (of all machines) to that of the entire bread output. The average investment period of the services of machines will depend on the ratio of the value of the entire machine-service input (of all slaves)[48] to that of the

[46] It would necessarily be true therefore of a slave state that both capital and labour yield a positive rate of return, irrespective of the extent of accumulation (unless there is some third "fixed" factor, like land, in relation to which both become less productive, by an increase in their quantity). But it will normally be true even in a nonslave state that the rate of return will be positive on both "machines" and labour (though the latter, owing to the inalienability of the ownership of labour, can only be calculated on rather arbitrary criteria) although, of course, there will no longer be forces operative which tend to make them *equal*. But the rate of return, on one or the other, could fall to zero in "extreme cases": (1) when the quantity of labour has increased, by multiplication, to the extent that the marginal productivity of labour has been brought down to the labourers' subsistence level (the "stationary state" of Ricardo and the classics); (2) when the quantity of material resources has increased, by accumulation, to the extent that the marginal productivity of the services of capital goods has been brought down to the level of their "maintenance costs" (the stationary state of Professor Schumpeter). There seems to be no reason to assume that in the real world forces are operative which will inevitably draw the system either to the one or the other "extreme" of stationariness.

[47] If this rate is such that people are willing to save at that rate (and this desire, in the absence of a change in psychology, could only be strengthened by continued accumulation) our society would resemble the "expanding universe"; it could never become stationary.

[48] The "machine-service input" of slave capital takes two different forms. (1)

entire bread output. Since both refer to the same bread output, an average between the two is completely meaningless. Both of these investment periods will, of course, remain unaffected by changes in the amount of capital.

If we now assume that, for some reason, the number of slaves is "held constant," when capital is accumulated, the increase in capital can only take the form of an increase in machines. Then the investment period of *labour* will rise, and the real rate of return on *machines* will fall. (Correspondingly, the investment period of machine services will *fall*, and the rate of return on slave labour will *rise*, but not to the same extent.) This "lengthening" of the investment period for slave labour can take various forms. (1) There might be an increase in the number of the *same* machines, and a substitution of machine services for labour services, in the production of bread; this will imply a reduction in the amount of co-operating labour, and an increase in the amount of invested labour, per unit of bread; (2) There might be an increase in the durability of machines, in which case the proportion of invested to co-operating labour can remain the same; (3) There might be a change in the "degree of automatism" of the machines (with or without a change in durability), in which case again the proportion of invested labour is increased and the proportion of co-operating labour reduced. All three cases imply a reduction in current labour input, and an increase in "initial input," per unit of bread output. If we now further assume that the slaves are liberated and in consequence only machines are regarded as "capital," the rate of interest will be determined by the yield of machines only; and we have then arrived at the Austrian theory of capital.

It follows from this analysis that the *Senior-Jevons-Böhm-Bawerkian law of roundaboutness is merely a roundabout way of expressing the law of nonproportional returns.* Once it is realised that the only difference between "produced" and "nonproduced" resources lies in the fact that the one can be augmented by economic disposition and the other cannot, it is clear that the ultimate reason why the rate of interest is falling with an increase in capital is precisely the same as the reason why rents are rising (or wages falling) with an increase in labour. A relative increase in the number of slaves, in the case where "land" and "slave labour" are the only scarce resources, could just as well be said to imply an increase in the "investment period" of the services of land, as a reduction of the marginal productivity of the services of labour; while the material content of the Austrian theory of capital

---

The services of machines directly co-operate with labour in producing bread. (2) Bread is also required for the maintenance of labour (which must be deducted from the "net output" of bread) and this maintenance bread also represents a certain quantity of machine service. (The same is true the other way round, of course.)

could be equally well expressed by saying that capital accumulation leads to a reduction in the marginal productivity of the services of those factors whose quantity can be augmented by such accumulation, as by saying that it increases the investment period of the services of those resources whose quantity remains constant.

The purpose of the "investment period" approach is to reduce the production function to two variables, substituting "waiting" for the services of all produced (or variable) factors, with interest as the price of "waiting." In this way—and only in this way—can *capital as capital* be treated as a factor of production, commensurate with "labour." This, however, can only be done so long as the services of the "fixed" factors can themselves be regarded as homogeneous, or at any rate sufficiently homogeneous to leave their relative scarcity unaffected by changes in the amount of the services of other resources. In the above example machine services and labour services were the only scarce factors. This enables us, by regarding the quantity of labour as constant, to measure changes in the amount of machine services available by changes in the "investment period" of the services of labour. Had we assumed three factors, say the services of machines, labour and land, among which only the services of machines could be increased in quantity by capital accumulation, neither the investment period of the services of land, nor the investment period of the services of labour would have afforded an unambiguous measure of the amount of machine capital. A combined "investment period" of the services or these "original" or rather constant resources, on the other hand, would have been possible only if the services of machines were assumed to be an "independent good" relatively to the services of land and labour, i.e., if the marginal productivity-ratio between land services and labour services depended only on the relative amounts of land service and labour service, but not on the quantity of machines.[49]

Further consideration shows, moreover, that the same objection which can be brought up as regards the nonhomogeneity of the services of fixed resources also applies as to the nonhomogeneity of final products. So far we have treated consumption goods—"bread"—as if they were a homogeneous entity, or if not homogeneous, at any rate something the composition of which can be regarded as given. It is obvious, however, that except in the special case where all consumption goods contain the services of fixed resources in the same proportions, an increase in the quantity of capital will lead to a change in the relative prices of different types of consumption goods, and thus to a change in the composition of the consumption stream. In that case it

[49] This defect of the Austrian capital theory was first pointed out by F. X. Weiss, "Produktions Umwege und Kapitalzins," *Zeitschrift für Volkswirtschaft und Sozialpolitik*, 1921.

will no longer be legitimate to speak of the degree of roundaboutness involved in producing a unit of "final output," since we no longer have an unambiguous measure of that unit. Nor can one ascertain (once allowance is made for the "circularity" in production) the degree of roundaboutness for each kind of consumption good, taken separately. For the contribution of the services of produced resources are diffused between different industries; and this renders it impossible to impute a definite proportion of the aggregate stream of "labour" to a single kind of consumption good.[50]

So far we have conducted our analysis under purely static assumptions, and found that even under these assumptions the investment-period concept leads into difficulties once allowance is made for the fact that both the relative prices of different kinds of labour (and land) and the relative prices of different kinds of consumption goods might change as a result of a change in the quantity of capital. It is not proposed here to examine the further difficulties that emerge once the static assumptions are, in one respect or another, relaxed; nor even to enquire how far the methods of "comparative statics" are legitimate for dealing with problems of capital accumulation. There can be no doubt that for an analysis of dynamic problems—and especially of the *par excellence* dynamic problem of the trade cycle—the investment-period concept could hardly be of any use. At the same time we hope that we have succeeded in demonstrating that the real objections against the "Austrian" capital theory relate to the *measurability* of the investment period, rather than to its *relevance*. It can be argued on many grounds (some of them emphasised by Knight, some already emphasised by earlier writers, such as Professor Fisher) that the "investment period" ceases to be a quantitatively measurable magnitude once one departs from the level of abstraction of Böhm-Bawerk's and Wicksell's writings. But this is a very different thing from maintaining—as Professor Knight maintained in various articles—that the investment-period concept is also wholly irrelevant, i.e., that even if conditions are postulated under which it can be measured, it will have no correlation with the quantity of capital and the rate of interest. In so far as it is possible to give an index to the "degree of roundaboutness," it can also be shown that an increase in capital, if associated with a lower interest rate, will necessarily imply the adoption of more roundabout processes.

NICHOLAS KALDOR

London School of Economics

[50] It is only in cases where (as in our world of houses) the input stream of each single capital good consists exclusively of labour, or where the services of all capital goods are completely specific (i.e., they only contribute to the production of one final good) that the "investment periods" for individual commodities can be *separately* evaluated.

ANNUAL SURVEY OF STATISTICAL TECHNIQUE: METHODS OF COMPUTING AND ELIMINATING CHANGING SEASONAL FLUCTUATIONS[1]

By Horst Mendershausen

I. FROM THE MEASUREMENT OF THE "TYPICAL" SEASONAL MOVEMENT TO THE MEASUREMENT OF ITS VARIATIONS

The seasonal movement of a series consists in a succession of rises and falls which repeat themselves year by year at nearly the same date, so that an annual periodicity appears distinctly and continually in the fluctuations of the phenomenon. This periodicity is in the last analysis to be ascribed to the annual meteorological cycle of the seasons, whose influence works through a number of social factors, e.g., the organization of society, the methods of production and transport, the habits of consumption, etc.

It is often assumed that the seasonal movement has a perfectly regular annual periodicity. This assumption, however, implies that not only the weather factors are constant for corresponding moments of every year, but also that the other influencing elements above mentioned do not change during the period studied.

This implication does not correspond to reality. Though very regular, the meteorological cycle is far from being perfect, whether measured by temperature or the quantity and frequency of rainfall, etc. On the other hand the social organization, the technique, habits, and policy of production, without being radically overthrown from one year to the next, present some instability, due not only to their evolution but also to the business cycle and other social events, like wars and labour disputes. Consequently the measurement of a phenomenon which is influenced by all these circumstances can hardly be based on the assumption that it recurs with perfect regularity year by year.

It is true that in many cases a good approximation can be obtained by measurements which are based on the assumption that the seasonal movement is perfectly recurrent. But although in such cases it may be taken that the typical characteristics have successfully been determined, a full measurement will have to consider not only these average features, but also the deviations from it.

While the methods dealing with a stable seasonal normal have been developed to a stage where nothing of importance remains to be done, the measurement of variable seasonal movements is still in its beginings.

[1] The author would like to thank Miss Margaret F. W. Joseph and Dr. Hans Staehle, Geneva, for many useful ideas which have arisen out of discussions he has had with them on the subject of this paper, and for their assistance in the preparation of the English text.

Since the early days of business-cycle analysis by more refined statistical methods, marked by Persons' pioneer investigations, the question has arisen, whether seasonal fluctuations must be considered as normally stable or normally changing. At the outset the first point of view predominated. W. M. Persons[2] did not consider a variable seasonal movement at all. He put every deviation of a gross seasonal value[3] from its stable normal on a par with an error of observation.[4] "A seasonal movement is, by definition, a regular fluctuation occuring each year. There is disagreement among statisticians as to the logical possibility of securing indexes of seasonal variations that *change* each year."[5] But since the contributions of E. C. Snow (1923) and W. I. King (1924) a more realistic way of looking at this problem has spread,[6] without, however, revolutionizing the practical procedure used in eliminating seasonal movements from time series.

There is a curious contradiction between the increasing acceptance of the view that the seasonal movement is subject to variations, and statistical practice. The standard methods (e.g., Persons', Bowley's, and others) are still exclusively based on the stable normal. These methods give much attention to removing all "extreme" values from the series of the gross seasonal values, in order to obtain the "typical" seasonal component for each particular month. They work—quite logically—with the notions of median, extended median, central mean, etc., in their attempt to fix the "typical" seasonal movement and they do not carry the analysis further once this average has been secured.

Apart from certain difficulties of a more material and transitory nature—such as the inconsistency of the time series available and the

[2] W. M. Persons, "Indices of Business Conditions," *Review of Economic Statistics*, January, 1919.

[3] We mean by *gross seasonal values* the differences or ratios calculated between the original data and the non-seasonal (and non-accidental) component of the series, e.g., its 12-months moving average. We distinguish *gross seasonal differences* and *gross seasonal indices* according to the choice of differences or ratios for the representation of the seasonal movement.

[4] W. M. Persons, "Korrelation in Zeitreihen," *Handbuch der mathematischen Statistik*, Leipzig-Berlin, 1930, p. 207.

[5] W. C. Clark in Persons and others, *The Problem of Business Forecasting*, Boston and New York, 1924, p. 108.

See also W. L. Hart, "The Method of Monthly Means for Determining Seasonal Variations," *Jour. Am. Stat. Assoc.*, Sept., 1922, p. 343.

[6] Some hesitation in the adoption of the principle of the variable seasonal normal has still prevailed for a long time. Donner, e.g., is opposed to a normal which shows *short-time* fluctuations, while evolutive normals seem quite reasonable to him. See O. Donner, "Die Saisonschwankungen als Problem der Konjunkturforschung," *Vierteljahrshefte zur Konjunturforschung*, Sonderheft 6, Berlin, 1928, p. 56.

shortness of the periods they cover—two essential reasons for this divergence between the theoretical and the practical treatment of the question may be found in the following two considerations:

(1) In general the variations of the seasonal movement, if we disregard cases where there is no clearly established seasonal movement at all, are not very important in comparison with its "typical" values. The latter consequently provide a fairly good approximation to the true seasonal movement.[7]

As a general statement of facts this is certainly exact; but frequently very considerable variations occur. Thus, the German Institute for Business Cycle Research found it impossible to eliminate by the usual methods the seasonal components from German economic series during the year 1929 in consequence of the repercussions of the exceptional cold during the first months of this year.[8] O. Donner[9] gave up the construction of a seasonal normal for the unemployment of building workers, because the variations of this seasonal movement due to variations of winter temperature were too violent. W. I. King[10] mentioned similar difficulties encountered by E. Sydenstricker and R. Britten in their analysis of the causes of influenza.[11] These quotations, to which others could easily be added, show that the seasonality frequently varies to an appreciable extent.

(2) Speaking of Sydenstricker's study, King stressed the interest of a seasonal index which takes the climatological conditions of every year into account. But he thought it impossible to obtain such an index because weather is not the only factor influencing the seasonal movement. "Manifestly," he said, "the ideal index would be one which reflected accurately every shift in weather conditions. To obtain such an index appears, however, to be entirely impossible, for one cannot tell whether an exceptional rise or fall in any one year is due to the weather or to some other cause."[12]

In this statement, King neglected the possibility of examining the influences of the most important factors of the variations by multiple correlation analysis.

It seems therefore that there is no sufficient reason to hesitate in

[7] W. C. Clark, *op. cit.*, p. 117.

[8] *Vierteljahrshefte zur Konjunkturforschung*, Part A, Aug. 21, 1929, p. 36.

[9] O. Donner, "Die Saisonschwankungen der wichtigsten Wirtschaftsvorgänge in Deutschland seit 1924," *Vierteljahrshefte zur Konjunkturforschung*, Sonderheft 11, Berlin, 1928, p. 30–31.

[10] W. I. King, "An Improved Method for Measuring the Seasonal Factor," *Jour. Am. Stat. Assoc.*, Sept., 1924, p. 304.

[11] According to King these authors were the first to devise a method of obtaining a variable seasonal normal.

[12] W. I. King, *ibid.*

the development of methods which will permit the measurement of variable seasonal movements.

II. MECHANICAL ANALYSIS OF THE VARIATIONS OR ANALYSIS OF THEIR CAUSALITY

Two important groups may be distinguished among the methods which achieve this purpose:

A. Methods which attempt to measure and to eliminate variable seasonal movements by purely mechanical procedures;

B. Methods which attempt to arrive at variable seasonal normals on the basis of a causal explanation; once one or more important causal factors have been determined, it is their variation which accounts for the variation in the seasonal movement.

Obviously, the general approach of the second group has greater value from a scientific viewpoint. It represents an essential advance over the first in that mere mechanical procedures are necessarily unable to offer more than provisional or apparent solutions: science has attained its object only when a phenomenon is explained in terms of cause and effect. This is, of course, no criticism of the earlier writers since scientific development always passes through the stage of mechanical attempts at solutions, and consists in a slow process of elimination of such methods.

A. Methods which Attempt to Measure and to Eliminate Variable Seasonal Movements by Purely Mechanical Procedures

The two above groups are represented by very unequal numbers of writers. The former comprises many attempts which show a great variety of formal characteristics, while the second is made up of a few attempts only. For this reason the methods of the first group have been subdivided according to formal criteria, as follows:

(1) Methods which allow for discontinuous changes in the seasonal movement;

(2) Methods which allow for an evolutionary tendency (a trend) of the seasonal movement;

(3) Methods which allow for short-run temporary changes of the seasonal movement.

(1) *Methods which allow for discontinuous changes of the seasonal movement.* Attempts of this kind have been made by the Harvard Committee on Economic Research[13] and by the London and Cambridge Economic Service.[14] They led to a subdivision of the total series

[13] *Review of Economic Statistics*, 1925, No. 4.

[14] A. L. Bowley and K. C. Smith, "Seasonal Variations in Finance, Prices, and Industry," *London and Cambridge Economic Service*, Special Memorandum No. 7, July, 1924, p. 15.

into several successive parts, to the establishment of a stable normal for each of these parts and finally to a comparison of these normals. From the methodological standpoint this procedure is not more interesting than any other determination of a stable normal. Its application must be limited to those cases where abrupt changes of the seasonal movement really occur in consequence of some abrupt transformation in the conditions of the phenomenon.

In the cases of evolutionary or transitory changes of the seasonality this method is insufficient. It has nevertheless frequently been applied to such cases with the inevitable result that the difference between the stable normals obtained is perfectly arbitrary and fictitious. S. Kuznets,[15] for example, in his most interesting book on seasonal variations in American industry and trade, makes large use of subdivisions yielding successive stable normals in order to characterize evolutionary transformations in seasonality (cf. his table XXXIII). In several cases the figures given in his book (e.g., his charts 50 and 51) permit of establishing that there exists in reality neither an abrupt nor an evolutionary change of seasonal amplitude but only transitory variations. Examples of cases, where the difference of two successive stable normals is fictitious are: receipts of sheep and lambs and of hogs at primary markets, raw wool consumption, and steel ingot production (pp. 284 ff.).

(2) *Methods allowing for an evolutionary tendency (a trend) of seasonality.* This group of methods may be divided into two sections:

(a) Some authors maintain the notion of a stable seasonal normal, but they enlarge this concept by defining the stability as a *constant ratio* between the seasonal differences and the secular trend values of the original series. They suppose the seasonal differences to be proportional to the trend and eliminate their variation simply by taking their ratios to the corresponding trend ordinate. The methods of W. M. Persons,[16] A. L. Bowley (methods: Bowley III and VI),[17] G. R. Davies,[18] R. E. Chaddock,[19] H. D. Falkner,[20] and L. W. Hall[21] must be mentioned under this heading.

Bowley III, Davies, Chaddock, Falkner, and Hall fit linear trends, Bowley VI and Persons exponential ones. Falkner and Hall eliminate

[15] S. Kuznets, *Seasonal Variations in Industry and Trade*, New York, 1933.

[16] W. M. Persons, *Indices of Business Conditions*, pp. 16 and 19 ff.

[17] Bowley and Smith, *op. cit.*

[18] G. R. Davies, *Introduction to Economic Statistics*, New York, 1922.

[19] R. E. Chaddock, *Principles and Methods of Statistics*, Boston, 1925.

[20] H. D. Falkner, "On the Measurement of Seasonal Variations," *Jour. Am. Stat. Assoc.*, June, 1924, pp. 167 ff.

[21] L. W. Hall, "Seasonal Variation as a Relative of Secular Trend," *Jour. Am. Stat. Assoc.*, June, 1924, pp. 156 ff.

the trend before calculating the gross seasonal values, while the other methods do so in the last stage, the trend being eliminated from some kind of average seasonal values. We do not go here into the details of these methods. A convenient survey has been made of them by O. Donner.[22]

(b) Other authors go straight on to a variable seasonal normal, as, e.g., E. C. Snow,[23] W. L. Crum,[24] F. L. Carmichael,[25] H. and F. Hotelling.[26] The shape of the trends they envisage is generally linear. They are no more fitted to the original data in their chronological order but to the series of gross seasonal values for each month. Snow fits these trends to the series of empirically given values for the same month in successive years, Crum to its link relatives or to the ratios of the empirically given values to the trend of the original series, and Carmichael to the first differences of the empirically given values. Each of these methods leads to 12 monthly trends, fitted to the data in a more or less developed stage.

In Hotelling's study we find trend parabolas of the second or higher degree, fitted to the values of the original series. Snow and Hotelling aim at removing by one single operation the trend as well as the progressively changing seasonal movement from the data.[27]

These methods may be utilized in cases where a trend is really present in the gross seasonal values. The whole question of trend elimination is herewith opened. Much care is necessary in finding out whether the evolutionary tendency is not only an apparent one. The methods must be rejected in every case where the variation has equal chances to be evolutionary or transitory (perhaps periodic), and even if it is clearly evolutionary we must still establish whether the evolutionary factor might not better be represented by another independent variable than by the "factor time," a fiction involved in the trend calculus.

[22] O. Donner, "Die Saisonschwankungen als Problem der Konjunkturforschung," *op. cit.*

[23] E. C. Snow, "Trade Forecasting and Prices," *Journal of the Royal Statistical Association of London*, March, 1923.

[24] W. L. Crum, "Progressive Variation in Seasonality," *Jour. Am. Stat. Assoc.*, March, 1925, pp. 48 ff.

[25] F. L. Carmichael, "Methods of Computing Seasonal Indexes, Constant and Progressive," *Jour. Am. Stat. Assoc.*, Sept., 1927, pp. 339 ff.

[26] H. and F. Hotelling, "Causes of Birth Rate Fluctuations," *Jour. Am. Stat. Assoc.*, June, 1931, pp. 135 ff.

[27] Professor Hotelling has suggested in a letter to the present writer that the method used by R. A. Fisher for the analysis of "The Influence of Rainfall on the Yield of Wheat at Rothamsted," *Philosophical Transactions of the Royal Society of London*, Vol. 213 B, might be usefully applied to the problem of measuring variable seasonal movements.

(3) *Methods allowing for short-term temporary changes of the seasonal movement.* Several methods have been elaborated with a view to measuring transitory changes of the seasonal movement. The methods dealt with in this group are based on the assumption (most frequently tacit) that the seasonal movement itself or its variations show one or other of the characteristics enumerated in the following subdivisions.

(a) The absolute deviations from the 12-months moving average for any given month may vary from one year to the next but their ratios to the 12-months moving average[28] of the original series are assumed to remain constant. The method Bowley V[29] and the method ascribed to F. R. Macaulay[30] aim like those mentioned in the previous section (2) (a) at maintaining a stable seasonal normal in that they give a special sense to the notion of stability. They conceive of it as a stable ratio between the seasonal deviations of the different months and the 12-months moving average which represents, *grosso modo*, the trend and the cyclical component of the series. The seasonal deviations are permitted to vary to the extent to which variations of the 12-months moving average occur, so that the proportion between the two elements remains constant. When such a proportionality really exists, these methods may be useful, though they do not reveal the size of this proportion. If not, they are not better than any other stable normal. The assumption on which they are based has no particular a priori probability so that this latter case will not infrequently occur.

The maintenance of such a scheme in cases where its basic conditions are not fulfilled gives rise to curious roundabout ways. This may be illustrated by the attempt made by O. Donner[31] to construct a seasonal normal having the above special kind of stability for unemployment in Germany from 1924 to 1928. Donner states that the seasonal movement of this series does not show the desired proportionality and he therefore looks for another series describing the same phenomenon which is more suitable for the method. He finds such a series by making himself an *ad hoc* estimate of the number of *employed* persons. Donner constructs the stable normal for this employment series, eliminates the seasonal fluctuations from it, and subtracts month by month the non-seasonal number of employed from the total of workers in order to obtain the

[28] We do not distinguish here simple or centred 12-months moving averages. The latter are more correct, but the difference of the two methods is not considerable.

[29] Bowley and Smith, *op. cit.*

[30] *Federal Reserve Bulletin*, Vol. 8, Dec., 1922, p. 1416.—Macaulay, however, does not accept the paternity ("Meeting on the Measurement of Seasonal Variations of May 22, 1925," *Jour. Am. Stat. Assoc.*, Nov., 1925, p. 434).

[31] O. Donner, "Die Saisonschwankungen der wichtigsten Wirtschaftsvorgänge," *op. cit.*, p. 23.

nonseasonal number of unemployed. This latter series is finally subtracted from the empirically given number of unemployed and the result is a "variable seasonal normal" of unemployment. It may be questioned whether the maintenance of the scheme is worth all this trouble.

(a′) The absolute deviations for corresponding months from a general average of the series may change from one year to the next in consequence alone of changes in the seasonal *pattern*, but the normal seasonal differences for the maximum and minimum values of the seasonal movement are constant. Such a case is considered by F. Divisia.[31a] He proposes to compute a stable normal (average), not for the same month in successive years, but for the corresponding maximum and minimum terms chosen by graphical inspection of the given data. The subtraction of these normal values from the corresponding original data furnishes a series of nonseasonal points. They are equal in number to the maxima and minima considered. The lines joining the nonseasonal points are finally smoothed freehand.

This approach is a very rough one. Its basic assumptions are so limiting and the role of arbitrary decisions so important that it seems difficult to agree with Professor Divisia in preferring such a seasonal elimination to that obtained by a 12-months moving average.

(b) The normal seasonal values of any given month may change in either direction and at variable rates from one year to the next, but the change is assumed to be a gradual one. Under this head falls a certain number of methods which consist in fitting moving averages, moving medians, or moving central averages, etc., to the gross seasonal values of the same month in the different years. The variations of the seasonal character of each month is assumed gradual enough to be realistically represented by such a smoothing of the gross seasonal values.

Different methods of smoothing various kinds of gross seasonal values have been suggested: W. I. King[32] takes 9-years moving medians of the ratios between the original values and their trends; O. Gressens[33] 5-years moving medians, while J. C. Clendenin[34] prefers

[31a] F. Divisia, "Note sur le calcul de la variation saisonnière," XXIII^e session de l'Institut International de Statistique, Athens, 1936.

This note came to my attention only after the proofs had been corrected. I am in complete agreement with Professor Divisia's introductory remarks on the scientific solution to be aimed at in treating the problem of changing seasonality, i.e., the necessity of its causal determination.

[32] W. I. King, *op. cit.*

[33] O. Gressens, "On the Measurement of Seasonal Variations," *Jour. Am. Stat. Assoc.*, June, 1925, pp. 203 ff.

[34] J. C. Clendenin, "Measurement of Variations in Seasonal Distribution," *Jour. Am. Stat. Assoc.*, June, 1927, pp. 213 ff.

freehand smoothing. H. M. Flinn[35] smooths the Personian link relatives by a 7-years moving central average, and A. Joy and W. Thomas[36] treat the ratios between the original values and their 12-months moving average by 5-years moving averages.[37] F. R. Macaulay[38] proposes an analogous but a little more complicated method. A first approximation of the variable seasonal normal is obtained by a manipulation similar to that advocated by Joy and Thomas, i.e., by a 9-years moving central average of the gross seasonal differences. The latter are calculated between the original values and those of a 43-term graduation. The definitive variable normal consists in the differences between the first approximation and its own 12-months moving average.

The utility of this group of methods depends on the presence or absence of the supposed gradual character of the variations in the seasonal movement.

(c) The seasonal movement may change in any way from one year to the next, but it is in itself assumed to be very gradual and to approach a sine curve. Under this head fall the methods due to O. Anderson and R. Frisch.

According to Anderson's[39] view the gross seasonal differences, calculated between the original series and its 12-months moving average may be conceived as being made up of two components: their mathematical expectations and accidental deviations. The genuine seasonal movement is formed by the series of the mathematical expectations. In order to obtain it the accidental deviations must be eliminated. Anderson smooths for this purpose the original series by weighted moving averages, whose length and weights must be determined by the "variate difference method." By this method the smoothing effect of the moving average is intended to be made such as not to mitigate the genuine seasonal fluctuations. The latter is removed by a 12-months moving average from the series corrected for accidental influences and is represented by the differences between this moving average and the corrected series.

Anderson's method postulates two formal features of the seasonal fluctuations:

(α) The typical seasonal movement, i.e., the succession of the

[35] "Meeting on the Measurement of Seasonal Variations," *op. cit.*

[36] A. Joy, and W. Thomas, "The Use of Moving Averages in the Measurement of Seasonal Variations," *Jour. Am. Stat. Assoc.*, Sept., 1928, pp. 241 ff.

[37] This method is used by the Federal Reserve Board and termed "moving" seasonal index."

[38] F. R. Macaulay, *The Smoothing of Time Series*, New York, 1931, p. 130.

[39] O. Anderson, *Die Korrelationsrechnung in der Konjunkturforschung*, Veröffentlichungen der Frankfurter Gesellschaft für Konjunkturforschung, Heft 4, Bonn, 1929. See especially pp. 77 ff.

mathematical expectations, must be very gradual and approach a sine curve. The more it distinguishes itself from a gradual movement, the more the elimination of "accidental" deviations removes parts of the seasonal fluctuations. A. Wald[40] has demonstrated that this danger is not only restricted to extreme cases where the seasonal movement consists of one large deviation in one sense and eleven small ones in the other (as, e.g., in the series of purchases of household articles in Austria), but even arises in cases where the seasonal movements are much more gradual. We believe that Wald is quite right in opposing the application of the "variate difference method" to original data and in limiting it to series from which the seasonal fluctuations have already been eliminated.[41] All seasonal influences must indeed be eliminated, before Anderson's method of eliminating accidental deviations can reasonably be applied.

(β) Furthermore, Anderson's method tacitly postulates that all variations of the seasonal movement affect the neighbouring original values in such a way that the curve of the seasonal differences does not become irregular. Indeed, all small peaks and troughs of the curve, though they may be due to an exceptionally cold winter or to similar essentially seasonal reasons, are systematically ironed out in the course of his elimination of accidental elements. This postulate therefore is in contradiction to reality. It is curious to see Anderson explicitly draw attention to the influence of such an obviously seasonal factor as winter temperature upon the prices of eggs in the United States and nevertheless eliminate the influence of its changes as an "accidental" one. Thus the mathematical expectations of the seasonal differences are not able to represent the true seasonal movement.

R. Frisch has developed the following method of calculating variable seasonal indices.[42]

He first considers the case of a monocephalic seasonal movement (one minimum and one maximum during the year). An unweighted moving average is selected such as to be about as long as the shortest distance between two estimated seasonal "normal points"[43] and not

[40] A. Wald, *Berechnung und Ausschaltung von Saisonschwankungen*, Beiträge zur Konjunkturforschung, Herausgegeben vom Oesterreichischen Institut für Konjunkturforschung, Vol. 9, Vienna, 1936, pp. 35 ff.

[41] It is difficult to see why he exempts price series from his judgment and so approves the procedure of G. Tintner, *Prices in the Trade Cycle*, Vienna, 1935.

[42] Professor Frisch was kind enough to communicate to the present writer a brief outline of his method. This has frequently been used in his statistical seminar at Yale University in 1930 and subsequently in Oslo, but so far has not been published.

[43] The definition of these "normal points" follows from the further steps involved in the method. See below.

so long as completely to annihilate the seasonal swings, e.g., a 5-months moving average. This moving average is applied to the original series and furnishes the smoothing I. This latter curve is again smoothed by a moving average of the same kind and length, smoothing II, which is again followed by smoothing III. Four distinct curves are thus obtained: the original curve and the three successive smoothings. If there are no strong cyclical fluctuations in the original series the three derived curves will intersect periodically at the same points: the "normal points." These points are very important. Frisch interpolates through them a curve which may be called the "normal curve." The interpolation may be done by any process, also graphically. Then the deviations of any one of the successive smoothings from the "normal curve" are read off. If a very smooth seasonal index is desired, it is the deviations of the third smoothing from the "normal curve" that should be taken. These deviations are assumed representative of the seasonal fluctuation, except for their amplitude which is *damped.* In order to derive the final values of the seasonal index the individual deviations must be "inflated" by multiplying them by an *inflation factor:*

$$\left(\frac{\sin \lambda m}{m \sin \lambda}\right)^3,$$

where $\lambda = \pi D/p$; m represents the number of terms in the unweighted moving average; p is the period of a sine curve, determined for each moment t as twice the time distance $(t_2 - t_1)$ between the points of time t_2 and t_1 where the normal points immediately before and after t are situated; and D stands for the time distance between the observations which must be, and usually, of course are, equidistant. Thus the inflation factor changes in time, from one normal point range to the next. The final index of seasonal variation is variable and represented by a curve running continuously through time.

In the case of bicephalic or polycephalic seasonal movements the method to be applied is analogous but somewhat more complicated.

Some essential features of Frisch's method are open to serious objections:

(α) With regard to the successive smoothings of the original curve and the representation of its seasonality by a very smooth sinusoidal fluctuation, the same criticism applies as above made against Anderson's method. Frisch's method runs an even greater risk of modifying the shape of the observed seasonal movement. It does not only even out all small peaks and troughs, but manipulates the movement in such a way as to make it similar to a sine curve. If the real seasonal movement is not indeed sinusoidal (which seems to be a very rare case

in practice), the seasonal obtained by this method is as irrealistic as it is smooth.

(β) The "normal curve" is based on a series of interpolations whose number is not only considerable but also too large in our opinion. For, if the three smoothings intersect really periodically in *one* point—which is not always the case, especially if cyclical fluctuations occcur—the intersection points coincide practically with points of a 12-months moving average, and this latter supplies us at the same time with numerous intermediate points. Why should these points be abandoned in favour of the values interpolated between the "normal points"? Certainly the 12-months moving average is no perfect representation of the cyclical plus the evolutive movement. It tends to damp the cyclical movement, to eliminate short-time cyclical fluctuations, and to displace turning points within the range of six months. But an interpolation between the normal points does not guard against such mistakes. It of course permits the introduction by freehand corrections of all particular movements which seem realistic to us. But if such freehand corrections are desired they can as well be applied to the 12-months moving averages which have the comparative advantage of easier computation.[44]

(d) The changes of the normal seasonal values of the particular months are assumed to be functions of changes in the seasonal movement of the whole year, as regards its amplitude and (or) its shape. The year is either a calendar or a "moving" year. This assumption has been made in some recent approaches to the measurement of seasonal movements changing in the short run, first by S. Kuznets,[45] than by G. Gräbner[46] and A. Wald.[47] The great attention which has been given to some of these methods makes a more detailed analysis worth while.

Kuznets starts from the necessity of establishing a changing seasonal normal. Since the gross seasonal values deviate from the stable normal, factors of correction must be looked for, allowing the normal to be brought into harmony with the seasonal swing of the original values

[44] Great caution is, of course, required in the use of 12-months moving averages in the case of changing seasonal fluctuations and strong cyclical variations. They should either be corrected, or replaced by the values obtained on the basis of an elimination of the changing seasonal fluctuations through allowance for the variation in their causal determinants.

[45] Kuznets, *Seasonal Variations in Industry and Trade, op. cit.;* Kuznets, "Seasonal Pattern and Seasonal Amplitude: Measurement of their Short-time Variations," *Jour. Am. Stat. Assoc.*, January, 1932, pp. 9 ff.

[46] G. Gräbner, "Der 'bewegliche' Saisonindex," *Allgemeines Statistisches Archiv*, Vol. 24, 1934, pp. 143 ff.

[47] Wald, *op. cit.*

during consecutive years. Kuznets proposes to modify the stable normal in such a way that its *amplitude* corresponds for each year to the total amplitude of the observed gross seasonal values. Provided the shape of the seasonal curve is always the same, i.e., that the seasonal upswing as well as the downswing occur in all years during the same months and the proportions between the gross seasonal values of the particular months remain constant, a simple transformation of the normal amplitude furnishes the desired fit. The result aimed at consists of the stable normal plus a series of corrective factors, one for each calendar year. The corrective factor is a coefficient of linear regression of the gross seasonal amplitude on the normal amplitude:

$$b = \frac{\sum ds}{\sum s^2}$$

and is termed "amplitude-ratio." The d's represent percentage deviations of the original data from the 12-months moving average and the s's deviations of the normal seasonal index from 100.

The "amplitude-ratio" equals $+1$, when the gross seasonal amplitude for the year considered coincides perfectly with the normal, so that every d equals its corresponding s. Always provided that the pattern of the seasonal movement is the same in all the years (and therefore also in the normal) $b > +1$ means a supernormal gross amplitude and $b < +1$ a subnormal one.

If the aim of the analysis is only a comparison of seasonal amplitudes for different years or between the gross amplitude of a given year and the normal one, the "amplitude-ratios" provide the definitive answer. But besides this, these measures may also be used for establishing a variable normal. On the tacit assumption that a certain value of b is also valid for each of the 12 particular pairs of d and s of a given year, Kuznets proposes to multiply each such s by b in order to obtain the normal for this year, and so on for every other year. Thus a special seasonal normal is obtained for each calendar year.

Gräbner endeavours to extend Kuznets' method to the analysis of modifications in the shape of the seasonal normal. Kuznets himself has proposed to measure the degree of coincidence between the normal and the shape of the gross seasonal values by a coefficient of correlation between the d's and s's:

$$r = \frac{\sum ds}{\sqrt{\sum s^2(\sum d^2 - 12c^2)}},$$

where the c's represent correction values, used for centring the 12 d's on 100. Gräbner utilizes this measure in order to fit the shape of the

stable normal to the shape of the gross seasonal curve of the year. He writes for the seasonal normal of a given month: $100+k_{d_s}$, where k_{d_s} is calculated by the formula:

$$k_{d_s} = s - (1 - r)\left(s - d + \frac{1}{12}\sum d\right).$$

If the phenomenon shows at the same time variations of amplitude and of shape, Gräbner first follows Kuznets in establishing seasonal values corrected for changes in amplitude (s') and then calculates r and the k_{d_s}'s on the base of the s'-values.[48] The values of $100+k_{d_s}$ are then used to eliminate the seasonal fluctuation.

The meaning of Gräbner's formula may be interpreted as follows: The variable part of the seasonal index (k_{d_s}) of a given month equals the difference between s (the stable normal) and $s-d$ (deviation of the gross seasonal value from its stable normal), the latter being multiplied by a factor which may vary between 0 and 2 according to the difference between r and $+1$.[49]

If $r=+1$, that is to say, if the proportion between d and s is the same for all the 12 months of the year: $k_{d_s}=s$. In the case of a perfectly constant shape of the seasonal movement, Gräbner's variable normal equals the corresponding value of the stable normal, whatever the difference between s and d may be.

If r diminishes and approaches 0, the factor $1-r$ becomes more and more important and $s-d$ exerts a growing influence on k_{d_s}. If $r=0$, $k_{d_s}=d$, that is to say the variable normal becomes equal to the corresponding gross seasonal value, whatever the stable normal (s) may be.

If r becomes negative and approaches -1, the influence of $s-d$ on k_{d_s} grows further. If $r=-1$, $k_{d_s}=2d-s$. Thus, if the gross seasonal values have signs opposite to those of the stable normal and are in some fixed proportion to its values, the variable normal will either equal $3d$ (in the case of the proportion: $d/s=-1$) or in general $2d+f(d)$, since $-s$ is in this case a linear function of $+d$.

Gräbner stresses his intention not simply to replace the stable normal (s) for certain months by the gross seasonal values (d), but to "reconstruct the average relationship" between the s's and the d's. It is true indeed that he does not replace s by d throughout, but it seems doubtful whether his method furnishes a more valuable result than would be obtained by such a procedure.

[48] The intelligibility of Gräbner's article is handicapped by several incorrect definitions and errors in the formulas (cf., e.g., p. 157 of the article).

[49] The small value of $1/12\sum d$ may be neglected.

Gräbner's method calls for serious criticism:

(α) If the seasonal movement is completely revolutionized and the correlation between gross and normal seasonal values becomes negative, Gräbner's method leads to "corrections" which are void of any sense. There is no reason why s should be replaced by $2d+f(d)$, rather than by any other function of d.

(β) If the correlation between gross and normal seasonal values becomes positive but low, Gräbner replaces the stable normal by a variable one whose only justification consists in its fit to the gross seasonal values, and the goodness of this fit grows with the absence of correlation between stable normal and gross seasonal values of the year. This justification, however, is insufficient. It supposes that a deviation between the stable normal and the gross seasonal value of a given month has more seasonal significance, the higher is the positive correlation between these two values during the whole year; and less seasonal significance the lower this positive correlation. Such a hypothesis seems very doubtful. The difference $s-d$ may represent a seasonal anomaly as well as an accidental deviation, whatever the average relation between the d's and the s's over the whole year. The degree of correlation between the gross and the normal seasonal movement over the year does not permit of judging the real character of the gross seasonal value of a certain month. Such a judgment is impossible without knowledge of the changes in the causal factors of the seasonal swing. One may have this knowledge and apply it, but then Gräbner's formula becomes superfluous; or one may not have it and then Gräbner provides nothing but a very artificial hypothesis.

One objection may be made against the methods of both Kuznets and Gräbner, with regard to the fact that their corrections of the stable normal are constant during the calendar year and may change abruptly between December and January. The analysed phenomena hardly justify such a procedure. It is difficult to see why the variations of the seasonal movement must be constant during a whole year. This seems to be a formal requirement without any real reason. And why should the calendar year represent the most favourable period for equal (or proportional) changes in the seasonal cycle? The calendar year frequently cuts up periods which are homogeneous from the standpoint of their seasonal situation, e.g., the winter season of the building industry, and which may be supposed to show similar changes in the seasonal influence. According to these methods one part of such a relatively homogeneous group of monthly values is separated from the other and corrected in a different way. Would it not be more reasonable to begin and to end the year—if a 12-months period is to be considered

at all as a convenient time unit—at a moment when the good seasonal situation normally supersedes the bad one and *vice versa* and when a transitory change in the seasonal movement is most likely to occur (e.g., for building, in spring and autumn)?

On this point Wald's method constitutes an improvement on Kuznets.[50] He seeks to obtain correction factors for each particular month of the series, by means of which the stable normal of the month can be brought into harmony with its gross seasonal value.

The nature of his "new method" consists in a combination of Kuznets' method with the well-known procedure of moving averages. This can clearly be seen from an analysis of the formula by means of which Wald calculates the value of the variable normal for a month k of the year i:

$$s_{ik} = a_k' \cdot \frac{\sum_{j=k-6}^{k+5} a_j' \psi_{ij}}{\sum_{j=k-6}^{k+5} (a_j')^2}.$$

In this formula a_k' represents the stable seasonal normal of the month k, e.g., the central arithmetic mean of the gross seasonal differences[51] for k within the range of years analysed (a_k), corrected so as to render the algebraic sum of the 12 a_k equal to 0. The stable normal (a_k') is multiplied by another factor (the right-hand part of the formula). If this latter equals $+1$, $s_{ik} = a_k'$, that is to say the variable normal equals the stable normal. If the factor diverges from $+1$, the two normals differ from each other. Now the factor on the right hand is a *coefficient of linear regression of the gross seasonal differences on the normal seasonal differences* of the year in the middle of which the month k is situated. This factor corresponds perfectly to Kuznets' "amplitude-ratio," if we replace the seasonal *differences* by seasonal *indices* and consider "moving" years instead of successive calendar years. The regression coefficient becomes equal to $+1$, if during the year considered each gross difference (ψ_{ij}) equals the corresponding normal difference (a_j'). It is $> +1$, when the mean product $a'\psi$ is greater than the mean

[50] Wald himself does not explain the relation of his method to those of his predecessors. He appears to consider it rather as a completely new approach.

[51] The replacement of the gross seasonal indices proposed by Kuznets and Gräbner, by gross seasonal differences is a priori neither good nor bad. The choice should depend on hypotheses concerning the nature of the seasonal phenomenon to be analysed. We cannot agree with the view that the validity of the notion of gross seasonal *indices* is bound up with the validity of the concept of the stable normal, an opinion which seems to be held by Wald (*op. cit.*, p. 78).

square a'^2; it is on the contrary $<+1$, when the relation between these two magnitudes is the reverse. If we suppose, with Wald, that the form of the seasonal curve remains constant for all years of the series, the first case will occur if each ψ is greater (in absolute value) than the corresponding a', and the second case, if each a' is greater than the corresponding ψ. Thus, provided that Wald's hypothesis regarding the constancy of the shape of the seasonal curve holds true, the regression factor expresses the relation between the gross and the normal seasonal amplitude. If, however, the hypothesis is not fulfilled, the value of the regression factor cannot be regarded as expressing the relation between the amplitudes, and its utilization for the purpose of correcting the amplitude of the stable normal is therefore not justified. In this case the factor becomes a mere average of various increases, or of increases and decreases of the gross seasonal differences with respect to their normals.

After having calculated the series of s_{ik}, Wald proceeds to eliminate the seasonal movement of the original data, by subtracting from each item the corresponding s_{ik}.[52] If a periodical season-like movement is still found in the deviations of the residual series from the 12-months moving average of the original data, Wald applies a mechanical correction of the s_{ik}'s. The aim of this correction is to eliminate all season-like movements from the residual series. If the correction is successful, Wald obtains finally a series of normal seasonal differences, by means of which the original series can be reduced approximately to its 12-months moving average. The differences between the latter and the series resulting from Wald's seasonal elimination are termed "residual fluctuations" and are regarded as representing the casual component. Wald considers the 12-months moving average as giving a satisfactory representation of the "trend and cycle" component. He makes no attempt to improve the measurement of this component but tries only to separate the seasonal and the casual component of the analysed series.

Wald's method is very similar to that of Kuznets. The only essential difference seems to consist in the fact that Wald does not suppose changes of the seasonal movement to take place suddenly between the end of one calendar year and the beginning of the next, but assumes the changes to be gradual and "slow," so that they can be neglected within any given range of 12 months.[53] Thus the "jumping" correction factor of Kuznets is replaced by a "sliding" one in the case of Wald.

[52] The s_{ik}'s have positive signs for the months of the seasonal upswing and negative signs for the months of the downswing.

[53] Wald, *op. cit.*, p. 85.

After having indicated some of the essential features of the methods of *Kuznets*, *Gräbner*, and *Wald*, we may turn to a consideration of their general common line of approach. Are they useful for measuring changes in the seasonal movement? Do they improve the possibility of eliminating variable seasonal movements?

Let us consider first their *measuring* capacity. Kuznets' "amplitude-ratio" and Wald's regression coefficient are intended to measure changes in seasonal amplitude. But we have seen that these measures work satisfactorily only in the case of a constant seasonal shape. To suppose such a case to hold everywhere is hardly a very realistic hypothesis. Where there is considerable variation in the amplitude, changes of pattern are also likely to occur. It would be necessary to prove first, that the displacements of the gross seasonal values of the analysed series during a year all bear the same relation to the normal values (Kuznets), or that the displacement for a particular month k may be considered as an average highly typical of the displacements observed for the monthly values from $k-6$ to $k+5$ (Wald). Even if such an hypothesis were possible in cases where the seasonal fluctuations were influenced during 12 months and more by a *constant* cause (e.g., structural changes, slow cyclical movements) it is certainly incompatible with any influences on the seasonal movement which are subject to rapid changes (e.g., temperature, rainfall, sudden cyclical movements). Thus the basic hypothesis underlying these methods seems to be a somewhat weak one.

Kuznets and Gräbner do indeed attempt to measure changes in the form of the seasonal curve by means of r. But this procedure too is based on very questionable assumptions. Every casual fluctuation is regarded as a deviation from the normal form. Gräbner admits that r has a sense only if there are *reasons* for the change of the seasonal pattern. But even if such reasons were known, it is hard to conceive what interest a calculation of r would offer so long as the *total absence* of random fluctuations had not been proved. Moreover, r is not insensitive to changes in the seasonal *amplitude* which do not affect the gross seasonal values of the particular months in the same proportion.

We may conclude that these measures are valid only in two extreme cases: (α) if the amplitude of the seasonal curve changes, but its form remains constant; (β) if the form changes either without accompanying changes in amplitude, or with exactly proportional changes in all the months of the year, and without any casual fluctuations. In other cases, neither the amplitude ratios of Kuznets and Wald, nor r, nor finally the two kinds of measures together express clearly the change which has occurred. It seems probable, that the great majority of practical cases will not be such that these measures may be justifiably applied.

The formal requirements for the validity of these measures correspond to the desire of their authors to find a purely mechanical mean of measuring changing seasonal movements. But if these authors would try to analyse the causes of the changes and the correlation between changing causes and changing seasonal movement they would surely realise the advantage of analysing changes of the *particular gross seasonal values* instead of those of amplitude and shape. The particular gross seasonal values, of course, represent elements of amplitude and form at the same time. After having determined the seasonal elements in these values, attention could be devoted to the question of changes in amplitude and form. Then and only then could the proposed coefficients of regression or correlation usefully be applied. All formal requirements of "constant form," "constant or proportionally inflated amplitude," "slow change," etc., could be dropped in such a determination of a variable seasonal normal.

From a practical standpoint the shortcomings of these methods reduce the significance of their results to such an extent that they offer no advantages over those obtained by much simpler procedures, as, for example, by simply measuring the gross seasonal values through a comparison of the empirical data with a 12-months moving average. Such a procedure would also provide us with a "variable normal." This would indeed be a very crude method, resting on highly questionable hypotheses (e.g., the pure nonseasonal and noncasual character of the 12-months moving average, absence of casual fluctuations in the gross seasonal values, etc.). But its results can hardly be regarded as less significant than those obtained by the more complicated of our authors, so long as these cannot give a clear, realistic sense to the *residual* series, composed of the differences between the 12-months moving average on the one hand, and the result of a seasonal elimination by means of their variable normals on the other. Such a test is not provided. Kuznets and Gräbner have not considered the residual series at all. Wald indeed considers it, but only to the extent of establishing that it is "residual." In correcting his seasonal measures (s_{ik}) to give a more irregular aspect to the residual series, he admits tacitly the weakness of his method. He does not prove that the final residual series is free from seasonal and certain cyclical influences, so that it might be considered as representing casual fluctuations only. And in fact this is impossible, once certain of the irrealistic working hypotheses regarding the nature of the seasonal movement and of the 12-months moving average have been dropped.

We thus come to the conclusion that the results of these methods in measuring variable seasonal movements are no more significant than those obtainable by a much simpler mechanical method; and that the

proposed measures of changes in amplitude or pattern would be useful only *after* the variable seasonal had been exactly determined. For purposes of a provisional analysis the simple device would seem to be quite sufficient. If, however, a truly scientific analysis is aimed at, any mechanical procedure must of course be excluded.

The shortcomings of these methods become apparent when they are used for the *elimination* of the seasonal movement. The elimination of a phenomenon depends on its measurement. It is not possible to eliminate anything accurately which has not been accurately measured. The above criticism must therefore be applied to the results of elimination as well as to those of measurement of the seasonal movement. In addition, these methods involve certain practical disadvantages as compared with the crude method of the 12-months moving average: (α) They imply more calculation work; (β) while the elimination by means of a centred 12-months moving average may be carried up to the 7th term before the last given value (and after the first given value) of the analysed series, the elimination by means of Wald's method cannot proceed further than to the 12th term before the last and after the first given value. The elimination by means of Kuznets' and Gräbner's method must stop at a distance from the beginning and the end of the series, which varies according to the position in time: If the original data necessary to compute the 12-month's moving average for a calendar year are just complete, the distance is at its minimum, i.e., 7 months, for the last December value and at its maximum (18 months) for the preceding January value. Thus the seasonal elimination cannot be extended to the extreme years, which often provide especial interest; moreover, extrapolation devices cannot furnish a substitute, without involving a further loss in significance.

For these reasons it seems that the methods here described represent no real progress in the field of eliminating changing seasonal movements.

While reading proofs, the writer made the acquaintance of a new method, proposed by R. Zaycoff.[53a] This author attempts to give a generalization of Wald's method:

(1) Trend and cycle are not only represented by the 12-months moving average of the original series, but by this average plus an expression allowing for changes in amplitude and pattern of the gross seasonal deviations. Thus the damping of the cyclical variation by the moving average can be corrected in a certain way.

[53a] R. Zaycoff, "Über die Zerlegung statistischer Zeitreihen in drei Komponenten," *Publications of the Statistical Institute for Economic Research, State University of Sofia*, 1936, No. 4. The article contains an application of Wald's and Zaycoff's methods on the series of wholesale prices for barley in Bulgaria.

(2) The seasonal component is constructed in such a way that it allows not only for "slow" changes in amplitude but also for changes in pattern. In calculating the stable seasonal normal Zaycoff tries to restrict the number of disregarded "extreme" values to an absolute minimum.

Zaycoff's working hypotheses seem somewhat more realistic than those of Wald, but the procedure remains mechanical and causal factors are only mentioned by the way, so that the real meaning of the seasonal and the casual component finally separated is rather obscure. The amount of computation work required by this method seems to be enormous, and numerous extrapolations are necessary to avoid a considerable loss of empirical data at the two ends of the analysed series; but Zaycoff as well as Wald proceed by a very thorough mathematical reasoning.

B. Methods which Attempt to Arrive at Variable Seasonal Normals on the Basis of a Causal Explanation

We come now to the last group of methods designed for the analysis of the causation of changing seasonal fluctuations.

Here again an attempt has been made to reconcile the concept of the stable seasonal normal with the unstable reality. H. B. Killough[54] calculates stable normals for subdivisions of the series. In this case it is a question of subdivisions not into successive periods, but into groups of years chosen according to some other principle. Killough, who analyses the seasonal movement of the price of oats, calculates 2 stable normals, one for years with a good crop, another for bad years. The remarks put forward under A (1) must be recalled here. To justify such a procedure it is necessary first that the two situations may be clearly distinguished and second that they were not susceptible of comparison by quantitative measurement. The first condition was given in Killough's case but not the second, since crops can be compared in respect to their quantity and to other measurable features. Therefore the two situations considered by Killough represent unnecessary simplifications. By maintaining the concept of a stable normal, Killough excluded the possibility of finding a more significant regression of the seasonal movement on the amount of the crop.

The methods of E. Gjermoe,[55] F. I. Zrzavy,[56] and J. Wisniewski concentrate directly on the variable seasonal normal. Gjermoe con-

[54] H. B. Killough, *What makes the price of oats?*, U. S. Department of Agriculture, Department Bulletin, No. 1351, Washington, 1925.

[55] E. Gjermoe, "The Seasonal Movements of Employment in their Relation to Business Cycle," *Nordic Statistical Journal*, Vol. 3, 1931, pp. 532 ff.

[56] F. I. Zrzavy, "Ausschaltung von Saisonschwankungen mittels Lag-Correlation," *Monatsberichte des Oesterreichischen Instituts für Konjunkturforschung*, Beilage n. 2, Vienna, Oct., 1933.

siders the relation between the variations of the seasonal amplitude during the year and the business cycle situation. Wisniewski chooses the same approach in his study published in ECONOMETRICA.[57] In his article on seasonal fluctuations in the building industry[58] he analyses the variations of the gross seasonal indices of particular months in relation to temperature, and recently he proposes an analytical process dealing with monthly gross seasonal values and the cyclical situation.[59] Zrzavy analyses the same relation.

Gjermoe supposes the cyclical situation of certain occupational series to exert, during a given year, an influence on the seasonal amplitude of these series. It may be assumed not without plausibility that the amplitude will be greater during a bad cyclical situation than during a good one.[60] On the supposition that the cyclical situation exercises a proportional influence on the seasonal deviation of all months of the year (see Kuznets), Gjermoe calculates the values of his dependent variable (s) in the following way: he establishes calendar year for calendar year the standard deviation (σ) of the gross seasonal differences from 100 and divides it by the average value of the 12-months moving average during the same year. This latter average value is also utilized as independent variable (g).

After eliminating a certain number of years, which show important changes in the level of the 12-months moving average or other disturbing influences, the author analyses the regression of s on g. Three regression diagrams show an interesting distribution of the points of coincidence. As g approaches its maximum (100 per cent), the relative seasonal amplitude (s) approaches its minimum (0). As g decreases, s increases, first rapidly, then more and more slowly, with a tendency finally to become constant. Gjermoe maintains (1) that the relative seasonal amplitude varies inversely with respect to the cyclical situation and (2) that this covariation is not linear but forms a branch of a hyperbola. Such a hyperbola appears indeed to fit the given points of coincidence rather well, so far as can be judged from a purely graphical inspection of the correlation. (The numerical value of the correlation is unfortunately not given by Gjermoe.) He indicates no reasons for his preference for the hyperbola, apart from the goodness of its fit.

[57] J. Wisniewski, "Interdependence of Cyclical and Seasonal Variation," ECONOMETRICA, April, 1934.

[58] J. Wisniewski, "Les fluctuations saisonnières dans l'industrie du bâtiment," (Polish text with French résumé), *Kwartalnik Statystyczny*, 1935, n. 2, pp. 239 ff.

[59] J. Wisniewski, "Note on Seasonal Variation," *Studia Ekonomiczne*, Vol. 3, Cracow, 1936.

[60] J. Åkerman, *Economic Progress and Economic Crisis*, London, 1932, pp. 41–42.

Kuznets, *Seasonal Variations in Industry and Trade*, *op. cit.*, pp. 336 ff.

Gjermoe, *op. cit.*, pp. 570 f.

Gjermoe concludes that the influence of the cyclical situation on the seasonal amplitude of the occupational series is well established; but that, because of the large deviations from the regression curve due to other influences, the measure of this influence by means of the regression hyperbolas cannot be relied upon. Gjermoe therefore does not propose to utilize his regressions to construct a variable normal.

The general method of approach of this analysis constitutes a definite advance. Gjermoe is no longer seeking a mechanical device for measuring changing seasonal fluctuations, but he constructs a bridge between the variations and one of their causes by means of correlation analysis accompanied by practical reasoning. The details of his study are nevertheless open to some criticism:

(1) The dependent variable (s) is too crude to give a true representation of the seasonal movement. We have already pointed out that it would be better to analyse the variations of the seasonal movement by considering the individual gross seasonal values than by taking only their amplitude during a whole year. The latter may fail to give sufficient representation of the changes of the particular values because of changes in the determining factor.

(2) Nor is the independent variable completely satisfactory. In the first place it seems doubtful whether the values of the 12-months moving average give an adequate picture of the cyclical situation. But even if this be taken for granted, it is again questionable whether yearly averages of these values are suitable for the measurement desired. The object is to pair each gross seasonal value with the cyclical index referring to some point of time in respect to which the lag is really significant for the supposed relationship. The fulfillment of this task is rendered especially difficult by Gjermoe's method of considering the seasonal amplitude of the year as a whole. In consequence he is obliged to exclude from his analysis years where sensible changes in the level of the 12-months moving average have occurred.

(3) Gjermoe draws the attention to the influence of other factors than the cyclical situation (such as temperature), but does not take them into account in the computation. This would moreover be difficult to do, because of the fact that he considers only the whole year's seasonal amplitude.

(4) The hyperbolic regression is chosen without reference to the condition of the phenomenon to be analysed. It is even inconsistent with some circumstances mentioned by Gjermoe in the text, e.g., the increase in the number of enterprises with low but regular employment in the course of a very accentuated depression, a fact which probably leads to a new reduction of the seasonal amplitude after its increase during the less acute stages of the crisis. In an analysis of similar statis-

tics, the present writer[61] arrived at the conclusion that instead of Gjermoe's hyperbolas, U-shaped regression curves must be looked for, for which realistic hypotheses with regard to the relationship considered could be put forward. Gjermoe's diagrams indeed reveal a tendency for the empirical regression to deviate from the hyperbola in a corresponding way, i.e., to approach a definite finite limiting value ($g=100$; $s=100$) (Gjermoe's diagrams 1, 2, and 3) and to begin a new downward movement in the left-hand part of the scale, so that a tendency towards an inverted U-shape appears (diagrams 1 and 3).

Zrzavy starts his analysis of the cyclical influence on the seasonal movement of unemployment by supposing—without any restriction—that its seasonal character is comparatively weak, if the level of cyclical unemployment (12-months moving average) is high, and strong if the latter is low. Persons' link-relative method applied to eliminate the seasonal movement furnishes fantastic results. The "correction for seasonal fluctuations" leads to a curve which shows surprising seasonal movements, parallel to the original one in good years (little unemployment), anti-parallel to the original one in bad years (high unemployment). In the first case Persons' stable seasonal normal eliminates too little, in the second case, too much. O. Anderson[62] has already shown by means of artificially constructed examples that this strange result might be expected from this method in certain cases of variable seasonal movements. This statement may be extended to cover all kinds of stable normals in cases involving a variable seasonal movement.

How does Zrzavy proceed in order to obtain a better seasonal normal?

(1) He establishes a 12-months moving average (or more exactly a centred 25-terms moving average of the half-monthly data) which enables him to calculate gross seasonal indices for every one of the 24 terms of each year.

(2) The moving average, and also the 24 series of gross seasonal indices of corresponding terms, show an upward trend during the 9 years (1924–1933). Zrzavy eliminates a linear[63] trend from all these series, in order to "simplify the procedure." He states that these trends do not represent a real evolutionary tendency but only "a magnitude in the computation" (Rechnungsgrösse).

(3) The next step is a calculation of simple correlation between the

[61] H. Mendershausen, "Les variations de la saisonnalité dans l'industrie de la construction," thesis of doctorate, to appear in 1937 at Geneva University.

[62] O. Anderson, *Zur Problematik der empirisch-statistischen Konjunkturforschung*, Veröffentlichungen der Frankfurter Gesellschaft für Konjunkturforschung, Heft 1, Bonn, 1929, p. 20.

[63] It is linear on logarithmic paper.

deviations of the gross seasonal indices for a certain term from their respective trend and the deviations of the values of the 12-months moving average from their trend, during the 9 years. Zrzavy computes lag-correlations between the two variables, by giving a lag of 2, $2\frac{1}{2}$, 3, . . . up to 7 months to the values of the 12-months moving average in relation to the particular term of the gross seasonal indices considered. Thus he gets for each term a series of lag correlations, from which he chooses the highest. The highest lag correlations for 10 terms are between 4 and 5 months. In these cases r is fairly high (>0.80). For 5 other terms the best lag is 2, 3, $3\frac{1}{2}$ or 7 months and for the last 9 cases Zrzavy does not state which is the most favourable lag. In these last cases (April, first half of May, October, November, and December) r is generally insignificant.

The sign of the correlation is not the same throughout the year. It is positive from the end of April until the end of November, and negative for the remaining parts of the year, particularly for the winter months. This means that during an unfavourable cyclical situation the seasonal indices of unemployment of the bad season (winter) diminish, those of the good season (summer) increase, so that the seasonal amplitude contracts. The favourable cyclical situation calls forth opposite tendencies resulting in an expansion of the seasonal amplitude.

Zrzavy affirms that a lag of $4\frac{1}{2}$ months between the two phenomena gives the best results. He proposes to utilize the regression coefficients (b) of these correlations in order to compute the variable seasonal normal (y) for every term:

$$y = a + bx.$$

a is the stable normal for the term considered during the 9 years, x the value of the 12-months moving average $4\frac{1}{2}$ months before; y, a, and x are expressed as percentages of the respective trends. Zrzavy utilizes this variable normal to eliminate the seasonal movement from the original series, and the result is better than that furnished by Persons' link-relative method.

As compared with Gjermoe's method, Zrzavy's procedure has the advantage of providing an analysis of the seasonal movement month for month. The cyclical influence is no longer considered as proportionate throughout all months of the year. Unfortunately Zrzavy does not give any coefficient of regression, so that it is impossible to say how the monthly regressions differ from each other. The differences between the correlation coefficients for the different terms suggest, however, that the relation between the variables does not remain constant. Moreover, Zrzavy's procedure seems open to serious criticism on a number of points:

(1) Zrzavy does not deal at all directly with the causal relationship between the cycle and the seasonal movement. He simply *postulates* a causal relation between them and then proceeds in very mechanical manner:

(a) There is no theoretical justification for the elimination of trends. It may "facilitate" the analysis, but that is hardly a sufficient reason. Zrzavy makes a "*reservatio mentalis*" with regard to the sense of such an elimination, but that does not provide it with a sense. By eliminating the trends he makes an arbitrary transformation of the original series for which no reasons can be found from the nature of the phenomenon in hand. The inclination of Zrzavy's "trend of unemployment" is, moreover, hardly plausible, unless we are prepared to believe that 20 years after 1933 more than half the Austrian population will be unemployed on relief. The further the proposed extrapolations of these trends are extended the larger are the divergences between the original data and the trend. It may be that these somewhat curious results account for the decision of the Austrian Institute of Trade Cycle Research not to content itself with this method. Within the scope of Zrzavy's own investigation the trend eliminations efface—and even reverse—the difference between the small crisis of 1925–26 and the severe one after 1929.

(b) Zrzavy gives no reasons in favour of the chosen lag of 4½ months; he shows that it furnishes the best results for 10 terms, but then proceeds to apply it to the computation of the normal values of all 24 terms.

(c) The question of a regression other than linear is not even broached.

(2) Zrzavy does not take into account meteorological influences on the seasonal movement, although their existence is revealed by the data. Thus his series,[64] corrected for changing seasonal fluctuations, shows in 1929 a seasonal movement, which probably has its origin in meteorological influences. During the particular cold winter of that year seasonal unemployment was higher than it should have been according to Zrzavy's normal; while during the following summer its reduction was greater as a result of the necessity to make up for the delayed activity.

(3) From a technical point of view, Zrzavy's study shows some weaknesses: (a) The correlation coefficients are not corrected for the number of degrees of freedom; (b) von Huhn's formula, by means of which Zrzavy compares the result of his elimination with that of Persons' method, is not well adapted for this purpose. The standard deviation is not a sufficient criterion in a case where the distribution

[64] Zrzavy, *op. cit.*, diagram, p. IX.

of the deviations *in time* is so important an element in the success or failure of a method; etc.

Wisniewski follows Gjermoe's method in his first study of the seasonal movement of employment in the Polish sugar industry.[65] The results are not very satisfactory. Meteorological and other influences reduce the goodness of fit of the (hyperbolic) regression of seasonal amplitude on the cyclical situation. With regard to the latter variable Wisniewski proposes, following a suggestion of Kalecki, to consider the relation between production and productive capacity instead of production only. He thinks that the value of this relation is of greater importance for the seasonal amplitude. It would seem worth while to attempt a practical verification of this interesting proposition.

Quite recently Wisniewski[66] has proposed another method of measuring variations of the seasonal movement in relation to the "normal" situation of the phenomenon. The "normal" situation is the result of evolutionary and cyclical factors. Wisniewski uses values of the 12-months moving average as representatives of normal situations (a). He computes a coefficient (P_i) which shows the linear regression of the original data of a certain month (x_i) during a series of years on the normal situation for the same month. Thus he obtains 12 regression coefficients (P_i), which permit him to calculate for every month of every year a normal seasonal difference:

$$w_i = P_i a + q_i - a,$$

where q_i is a constant value for the series of the month i, indicating the value of x on the regression line at the point where a equals 0. The values of P_i and of q_i are corrected in order to render $\sum_1^{12} P_i = 12$ and $\sum_1^{12} q_i = 0$.

If we disregard this correction, Wisniewski's method of measuring "linear seasonality" shows a great similarity to that of Zrzavy outlined above. Wisniewski writes his formula also as follows:

$$w_i = p_i a + q_i,$$

where $p_i = P_i - 1$ and represents the coefficient of linear regression of the gross seasonal differences of the month i ($x_i - a$) on the corresponding "normal situations" (a). This formula differs from Zrzavy's only by the following features:

(a) Zrzavy analyses gross seasonal indices, Wisniewski gross seasonal differences;

[65] Wisniewski, "Interdependence of Cyclical and Seasonal Variation," *op. cit.*

[66] Wisniewski, "Note on Seasonal Variation," *op. cit.*

(b) Zrzavy correlates deviations from an average (the trend), Wisniewski absolute values of the variables;

(c) For Zrzavy seasonality is a linear function of the cyclical situation, for Wisniewski of the "trend plus cycle" situation;

(d) Zrzavy takes into account a lag between the two variables, Wisniewski does not do so.

As compared with Zrzavy's approach, Wisniewski's method possesses both advantages and disadvantages, but the significance of the most important of the differences (points (a), (c), and (d)) cannot be judged without reference to the nature of the seasonal movement studied. In general, we may refer to several objections against Zrzavy's method (see above under (1) analysis of causal relationship, (1c) regression shape, and (2) meteorological influences).

In 1933, before the two above-mentioned investigations, Wisniewski published a detailed article on seasonal fluctuations in the building industry,[67] containing very interesting attempts at measuring meteorological influences on changes in the seasonal movement. Although this type of causal determination would seem to be the most obvious one, the above work was the first to attempt a statistical treatment of it. Wisniewski intended to measure the influence of temperature on the seasonal movement of unemployment in the building industry, but did not propose to utilize his measures for the immediate construction of a variable seasonal normal.

One of the questions he puts is: Is it possible to establish by statistical means the influence of temperature on the gross seasonal values, or even on the original data, of employment during certain months of the year?

Wisniewski establishes correlations between the percentages of employed (the original data, or the data divided by corresponding values of the 12-months moving average or by some other standard value) and the temperature (of the last month, or some weeks before the establishment of the employment figure). He assumes that this meteorological factor is more responsible for seasonal difficulties in the building industry than is rainfall, snowfall, etc. The correlations are in general not very good. They are disturbed by the intervention of other influences, mainly the business cycle; see, for example, the large divergence between observed and theoretical values for the years following the crisis of 1929 in Wisniewski's diagram on p. 282. Unfortunately Wisniewski does not take such influences into account in his computations. But he admits quite frankly that this should be done.

In connection with the structure of the regressions employed some reservations must be made. It seems probable that two different influences

[67] Wisniewski, "Les fluctuations saisonnières. . . .," *op. cit.*

of temperature on the seasonal movement must be distinguished, one consisting of a traditional reaction, the other of a reaction on the concrete conditions of the moment. In some of Wisniewski's regressions (e.g., for Poland) these influences are not separated and therefore a danger of a fictitious regression coefficient arises.[68] Finally in various cases the dependent variable does not represent only the seasonal (and accidental) portions of the data, but contains certain cyclical elements too.

Wisniewski's regressions are not very well constructed but constitute an interesting attempt to approach the causality of changing seasonal fluctuations. His studies, like those of Gjermoe and of Zrzavy, are tending in the right direction, i.e., to measure the changes of the seasonal movement as determined by changes in their causes. It would seem that a *general method* of dealing with this problem must proceed along these lines. It would involve a considerable development and generalization of the approach of these authors.

HORST MENDERSHAUSEN

Geneva

[68] Mendershausen, *op. cit.*, Chap. I.—In this study an attempt has been made to analyse the changes of seasonal movements in the construction industry of 5 countries in connection with changes in the cyclical situation, temperature and rainfall or snowfall, and to obtain a variable seasonal normal on the basis of the regressions.

SOME INVESTIGATIONS IN THE SAMPLING AND DISTRIBUTION OF RETAIL PRICES[1]

By John H. Cover

I. SAMPLING PROBLEMS

The basic problem in making a partial survey of retail prices is, of course, that of obtaining the most accurate and reliable representation of the entire retail price structure. If the objective is to get an average price for the community as a whole, that representative sampling method should be sought which gives a mean price with the smallest standard error. To solve this problem it is necessary to learn the manner in which the most probable price differs among localities.

Therefore, the first problem may be stated: Given a difference between two averages, what is the probability that such a discrepancy may have arisen through random sampling fluctuations? If the average prices of a commodity differ significantly between, say, chain and independent stores, the most reliable average of all prices is to be found, not in a random sample from all outlets, but in a "stratified" sample which contains representative prices from each homogeneous subgroup.

A second method of studying the retail price structure is an approach through the analysis of price frequency distributions. This permits a test of the validity of formulas such as that for the probable error. Moreover, bi-modality will sometimes provide evidence regarding the homogeneity of the universe; and it is often possible to dissolve bi-modal distributions into uni-modal series on the basis of kind of stores, type of retail operation, race or income of consumers, or other factors.

And thirdly, studies of correlation between price and factors (such as nationality, income) which may affect price have sometimes suggested significantly different price groupings. Through these techniques, many new lines of stratification other than geographical (which has always been applied unquestioningly) are found.

Problems to be considered here are ways and means of reducing the standard error, the accuracy of the standard error of an individual price, and accuracy of the standard error of a combination of prices. Some of the results of the survey and analysis concerning variations in price due to type of outlet, consumer groups, and geographical location have been published earlier.[2] These will be mentioned only in so far as they clarify the technique used in finding the most reliable average price and its standard error.

[1] The following collaborators are particularly to be credited with participation in this study: James Coble, Solomon Lischinsky, Henry Malcheski, Wilson Sweeney, and Harry Wilson.

[2] *Retail Price Behavior*, The University of Chicago Press, 1935.

Where significant differences are found among prices in different subgroups of the population, the variance of the mean of a sample may be reduced by an amount equal to the weighted mean square of the deviations of the subgroup means about the mean of the universe. Analysis of the data offers evidence supporting or disproving various hypotheses regarding the heterogeneity of the price universes.

The following are some of the characteristics of retail prices that make sampling analyses difficult:

1. The sampling universe is heterogeneous, and the probability of obtaining a given price varies with locality, class of store, and consumer group.

2. The price frequency distributions are not normal, being mostly very skew, leptokurtic, and frequently J-shaped or N-shaped.[3] This not only raises difficulties concerning standard errors of means but it also vitiates conclusions regarding the differences between average prices in different sections.

3. These frequency distributions vary greatly among commodities; and, therefore, a procedure which is devised for one commodity is not necessarily applicable to another.

4. The distributions are subject to change in time without notice.

II. STRATIFICATION

A few examples will illustrate the procedure followed in isolating homogeneous subgroups. In the first place, there was an attempt to break down drug prices (which very frequently exhibit striking bi-modality) into uni-modal distributions, i.e., stores were separated on the basis of kind of outlet, type of operation, nationality of customers, and so forth. In most cases the distributions for these subdivisions remained bi-modal, though occasionally an opportunity was provided for distinguishing between chain and independent stores, when their two uni-modal distributions combined to make a bi-modal. Therefore, it was necessary to accept the fact that unusual distributions were involved.

The next step was to find the evidence that favors a geographical stratification. Since the absolute difference between prices is not conclusive proof of underlying relationships, the standard error of the difference was computed in each case. Some error in the method was at times unavoidable because so much about the distribution of the differences of means of samples drawn from such skew universes remains unknown.

A good example of a situation in which the significances of price

[3] N-shaped is used throughout to indicate the special case of a J-shaped type in which there is a second mode near the center of the range.

differences are masked by high probable errors of these differences is found in drug prices in New York. In Brooklyn-Queens the difference between the average price in the White moderate-rental group and that in the Negro moderate-rental group was one cent, and the chances are that in 18 cases out of 100 as large or larger deviation would be found merely from sampling fluctuations. But in the Bronx where the price difference between the White moderate-rental group and a White lower-rental group is also one cent, the chances are 2 out of 100, which indicates a fairly significant difference.

The results indicate that geographical stratification is desirable. An unweighted arithmetic mean of twelve drug prices (drugs that are nationally advertised) varies as much as ten per cent between Minneapolis and Atlanta cash independent stores, while a standardized product such as Palmolive soap differs 33 per cent in price between the same cities. However, when prices are combined in larger groups and weighted, this extreme variation is not so apparent. For example, a preliminary study of data recently gathered indicates that the price index of a cost-of-living budget does not fluctuate to such a degree, and that the variations from city to city in food and clothing costs (two major items in the budget) are not highly correlated.

Hypotheses concerning the differences in prices among the various types of outlets were tested.

In St. Paul, a survey of food prices with complete coverage of stores offered an excellent opportunity for comparison of prices of fifty-eight selected food items among independent, co-operative, and sectional-national chain stores. Upon obtaining an arithmetic mean of the percentage differences in price between chain and co-operative stores for fifty-eight commodities, it was found that chains were about 8 per cent lower than co-operatives with independents halfway between them. These differences are clearly significant because the samples of price quotations were the total universes; therefore, this was the most conclusive result. Unfortunately, it cannot be said that this same condition necessarily applies to other cities, but it is important evidence.

To carry the comparison between chain and independent stores further, a study was made of the prices of five standard drug prescriptions in cash and credit chain and independent stores in four cities. The mean price differed between the lowest and highest of these groups—cash chain and credit independent outlets in New York—by 34 per cent. In three of the cities, cash independent stores were about 15 per cent higher in price than cash chain stores, while prices were about the same in Atlanta. No significant difference was apparent in Atlanta as between cash and credit outlets.

Since it was apparent that chain drugstore prices were definitely

lower than independent store prices, there was an inducement to study the factor of competition as a possible criterion for grouping neighborhood independent stores. Testing this hypothesis with twenty-one drugs, it was found that in only one case was the arithmetic mean of category A (independent stores *not* subject to chain-store competition) less than the mean of category B (independent stores subject to chain-store competition). The most marked difference appeared in the case of Palmolive soap where the mean of A exceeded the mean of B by 15 per cent. The geometric mean of all twenty-one A relatives with B as a base is 105.4 per cent. It is interesting to note the differences in price between independents which are subject to chain-store competition and the chain stores themselves. With the latter as a base, the geometric mean of B relatives is 108.7. It follows that the geometric mean of A relatives, with chain price as the base, averages 114.5 per cent. Therefore, prices appear to be significantly lower in trade centers in which there is competition between chain and independent outlets.

Though customer income (roughly approximated by rental groups) did not seem to play a part in the determination of drug prices, it was considered desirable to test this factor in the case of foods. A correlation table was set up to determine the relationship between rental groups and the prices of twenty-seven food items in Minneapolis; as far as possible, other price-determining factors that might vary with rental were held constant. The coefficient of correlation equals 0.48. With a group of eleven commodities this coefficient was raised to 0.65, and there was a correlation ratio of over 0.7. From this it might be concluded that there probably is some relationship between income and retail prices of food in Minneapolis. Whether this relationship will persist over a period of time, and whether rental is therefore a worthy criterion for stratification remains to be proved.

Several studies were made to determine the advisability of dividing the universe to be sampled into classifications according to race and creed. Drug and clothing prices in Negro neighborhoods of New York City were compared with those in White neighborhoods, and food prices were compared as between Kosher and non-Kosher stores in Minneapolis-St. Paul. After the rental factor was eliminated there was no conclusive evidence of relationships between these factors and prices. However, this subject deserves a wider study.

To illustrate the reduction in the standard error of the sample mean as a result of stratification, indexes of weighted sums of means of samples of six and of fifteen prices were calculated. The stratification was made on the basis of chain, neighborhood, and central shopping independents in Manhattan. For every commodity, the standard error of the mean of the stratified sample is less than that of the unstratified.

The probabilities that the sample index will differ from the true index by various percentages were worked out for stratified and unstratified samples and the differences between these probabilities were expressed as percentages of the probabilities for unstratified samples. The results are given in Table 1; these measure the gain in precision as a result of stratification. Thus the probability that an index made up of samples of six will deviate from the true value by more than 2 per cent, in February, 1934, is reduced 9.38 per cent as a result of stratification; and the probabilities of 5 per cent deviations are reduced about three and one-half times as much as the corresponding probabilities for deviations of 2 per cent.

TABLE 1

REDUCTION DUE TO STRATIFICATION OF THE SAMPLE IN THE PROBABILITY THAT AN INDEX WILL DIFFER FROM THE TRUE INDEX BY MORE THAN TWO AND FIVE PER CENT*

Data: Weighted sums of means of samples of six and fifteen prices quoted by chains, neighborhood independents, and central shopping independents in Manhattan, for six drug items.

February	Reduction in the Probability of Deviations Exceeding			
	2 Per Cent		5 Per Cent	
	Sample of 6	Sample of 15	Sample of 6	Sample of 15
1933	5.32%	11.41%	17.91%	43.28%
1934	9.38%	19.05%	32.91%	62.70%

* Reduction expressed as a percentage of the unstratified sample probabilities.

III. STANDARD ERRORS OF AVERAGE PRICES

Having approached the problem upon the assumption that the universe is not known precisely, this part of the analysis is a study of conclusions that may be drawn from definite knowledge of certain distributions.

Judging from large samples, the price distributions are generally skew (with various amounts of kurtosis), J-shaped, or multi-modal. In some cases, where complete universes have been obtained, our information is exact. For such universes, the distributions of means of small samples show different amounts of skewness, depending upon the universe and the size of sample; and the standard error, used with an assumption of normality, is not strictly applicable for small samples. In order to determine the amount of error involved in using these standard errors with average prices and price indexes, it is advisable to determine the distributions of means of samples drawn from such

non-normal universes and the size of sample necessary for given levels of accuracy. The well-known formula for the standard error of means of samples, $\sigma/\sqrt{N}$, is not as applicable here as is the corrected expression

$$\sigma_M = \frac{\sigma}{\sqrt{N}}\left(1 - \frac{N-1}{n-1}\right)^{1/2},$$

where σ is the standard deviation of the universe, n the total frequency, and N the number in the sample. This applies to finite populations (i.e., drawings without replacements).[4]

It is a familiar proposition that the distribution of the means of samples of N drawings tends toward normality as N is increased even when the universe is skew, but the aspect of greatest interest here is the rapidity with which normality is approached when the original universe is as unorthodox as those found among retail prices. The answer to this will give an estimate of the size of sample that is desirable and a measure of the error in the standard error assigned to the mean. Since the price distributions vary so widely, the approach to the problem has been to find the extreme cases and thus get outside limits of the error involved.

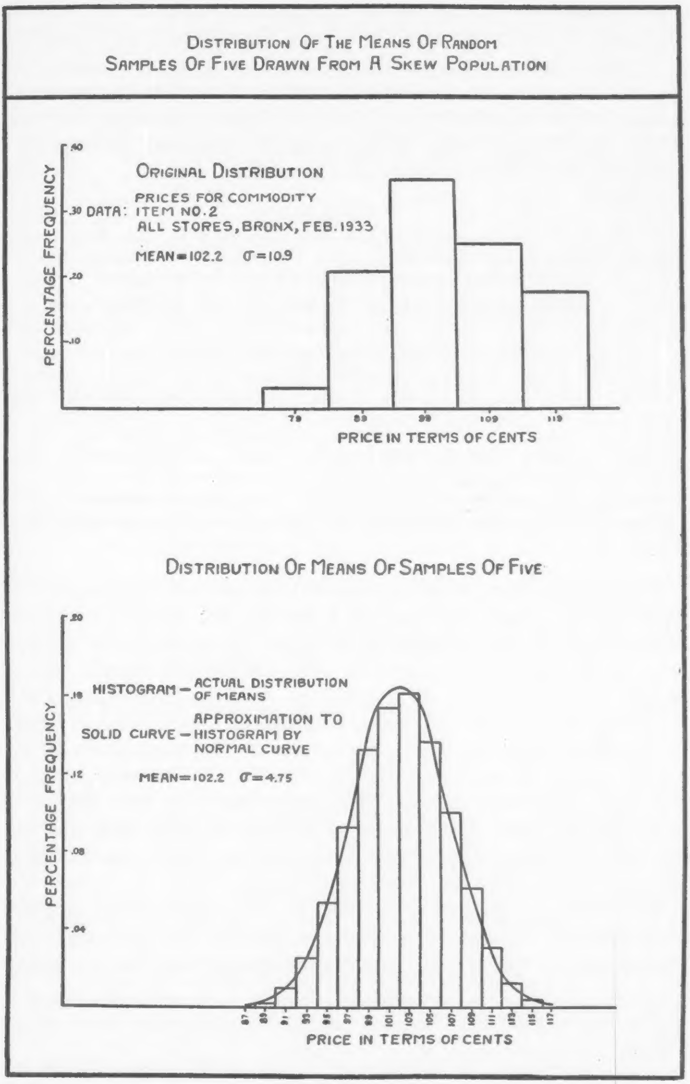
A moderately skewed, an N-shaped, and two J-shaped frequency distributions of drug prices in New York City were selected as distribution types. By the combinatorial method[5] the theoretical distributions of the means for different sized samples were worked out. For the slightly skew population, the distribution of means of samples of five only was calculated, because it was found that this approached normality fairly closely. A normal curve was fitted to it (Figure 1) and the mean devia-

[4] The formula for finite populations in many cases was found to give results greatly different from the formula for infinite populations. It is the standard error of the average price for that particular time; underlying relationships do not remain constant long enough to permit the calculation of an accurate standard error for a longer period of time.

[5] This method involves the following steps: All possible combinations of original prices taken N at a time (N being the number in the sample) are written down. The expected frequencies for each combination are calculated; and those frequencies whose combinations yield the same average price are added. When the frequencies for each average price are divided by the number of combinations of n things taken N at a time (where n is the number in the sampled population), the proportional frequencies are obtained.

For example, if prices .10, .20, .30, and .40 have frequencies in the sampled population of 3, 5, 10, and 4 respectively, an average price of .20 may be obtained for a sample of three in two ways; i.e., .20, .20, .20, and .10, .20, .30. Then the proportional frequency expected at this price is

$$(C_3^5 + C_1^3 \cdot C_1^5 \cdot C_1^{10})/C_3^{22}.$$



DISTRIBUTION OF THE MEANS OF RANDOM
SAMPLES OF FIVE DRAWN FROM A SKEW POPULATION
ORIGINAL DISTRIBUTION
DATA: PRICES FOR COMMODITY
ITEM NO. 2
ALL STORES, BRONX, FEB. 1933
MEAN = 102.2 σ = 10.9
PERCENTAGE FREQUENCY
PRICE IN TERMS OF CENTS
DISTRIBUTION OF MEANS OF SAMPLES OF FIVE
HISTOGRAM — ACTUAL DISTRIBUTION OF MEANS
SOLID CURVE — APPROXIMATION TO HISTOGRAM BY NORMAL CURVE
MEAN = 102.2 σ = 4.75
PERCENTAGE FREQUENCY
PRICE IN TERMS OF CENTS

FIGURE 1.

tion between the two curves was calculated to be 0.001714,[6] whereas the mean deviation between the original distribution and normality is 0.03725. This is a comparatively rapid approach to normality. One can now calculate the errors (in the probabilities that the mean of a sample of five will differ by two, five, and ten cents from the true mean of this distribution) which obtain when the normality assumption is made (Table 2). These errors are comparatively small.

TABLE 2

EXACT PROBABILITIES THAT THE MEAN OF A RANDOM SAMPLE OF FIVE FROM A SKEW UNIVERSE WILL DIFFER FROM THE TRUE MEAN BY MORE THAN 2, 5, AND 10 CENTS COMPARED WITH ESTIMATED PROBABILITIES*

Data: Prices for Drug #2, All Stores, Bronx, February, 1933

Error in Estimated Probability that Mean of Sample will fall								
Below M −2¢	Above M +2¢	Outside M ±2¢	Below M−5¢	Above M +5¢	Outside M ±5¢	Below M −10¢	Above M +10¢	Outside M ±10¢
.000	−.002	−.002	−.004	−.004	−.008	+.001	+.001	+.002

* Estimated probability is calculated on the assumption that means of samples are distributed normally. Minus figures indicate estimated probabilities which are too low.

The same procedure was carried out for an N-shaped population with samples of three, four, and five. For this population, assumption of normality for the distribution of means introduces a considerable error in the estimate of standard error, particularly for the smaller samples, which is due chiefly to the skewness of the distribution. Table 3, which consists of the differences between the probabilities of certain deviations as given by the exact distribution and those given by the normal curve, indicates that, with few exceptions, the error decreases with increasing size of sample.[7] The mean deviations from normality of the distributions of the means of samples of three, four, and five give other measures of the improvement with increase in size of sample. They are 0.070, 0.041, and 0.008, respectively.

The third population to be so tested was the J-shaped; this provides an estimate of the greatest error to be found in the standard error. For example, if one first assumes normality and says that the proba-

[6] A Chi-square test was not applied in this case because the discrepancy is that between two theoretical curves, in which case a Chi-square test is not applicable.

[7] Had the area under the normal curve been found to correspond to the discrete weights of the true distribution, the difference would have been much less and the decrease in error with increasing size of samples would have been uniform.

TABLE 3

EXACT PROBABILITIES THAT THE MEANS OF RANDOM SAMPLES OF THREE, FOUR, AND FIVE FROM AN N-SHAPED UNIVERSE WILL DIFFER FROM THE TRUE MEAN BY MORE THAN 0.5, 1, 2, AND 3 CENTS COMPARED WITH ESTIMATED PROBABILITIES*

Data: Prices for Drug #24, Independent Stores, Manhattan, February, 1934

Size of Sample	Error in Estimated Probability that Mean of Sample will fall					
	Below M −0.5¢	Above M +0.5¢	Outside M ±0.5¢	Below M −1¢	Above M +1¢	Outside M ±1¢
3	0.0761	−0.0661	0.0100	−0.0557	0.0538	−0.0019
4	0.0759	−0.1055	−0.0296	0.0334	−0.0905	−0.0571
5	0.0589	−0.0074	0.0515	−0.0014	−0.0189	−0.0203
	Below M −2¢	Above M +2¢	Outside M ±2¢	Below M −3¢	Above M +3¢	Outside M ±3¢
3	−0.0118	0.0568	0.0450	−0.0217	0.0088	−0.0129
4	−0.0039	0.0314	0.0275	−0.0034	0.0029	−0.0005
5	−0.0220	0.0194	−0.0026	−0.0001	0.0010	0.0009

* Estimated probability is calculated on the assumption that the means of samples are distributed normally.

bility is P that the mean observed in a sample will not differ from the true mean price by more than two cents, and if this probability is decreased by a correction factor obtained from a J-shaped universe, then

TABLE 4

EXACT PROBABILITIES THAT THE MEAN OF A RANDOM SAMPLE OF FIVE FROM A J-SHAPED UNIVERSE WILL DIFFER FROM THE TRUE MEAN BY MORE THAN 0.5, 1, 2, AND 3 CENTS COMPARED WITH ESTIMATED PROBABILITIES*

Data: Prices for Drug #19, Independent Stores, Bronx, February, 1934

Error in Estimated Probability that Mean of Sample will fall					
Below M −0.5¢	Above M +0.5¢	Outside M ±0.5¢	Below M −1¢	Above M +1¢	Outside M ±1¢
−0.0644	−0.0483	−0.1027	0.0057	−0.0059	−0.0002
Below M −2¢	Above M +2¢	Outside M ±2¢	Below M −3¢	Above M +3¢	Outside M ±3¢
−0.0120	0.0051	−0.0069	−0.0004	0.0001	−0.0003

* Estimated probability is calculated on the assumption that the means of samples are distributed normally.

one can be fairly certain that the reliability of the average has not been over-estimated. (Or the correction factor could be applied in the form of an increased number of items in the sample.) The differences in the probabilities as given by the exact histogram and the normal curve are great in the case of samples of five (Table 4); for deviations of a half cent, they are about 20 per cent. When the sign of the error is immaterial, the difference is considerably less.

TABLE 5

EXACT PROBABILITIES OF DEVIATIONS OF MEANS OF RANDOM SAMPLES OF FIVE FROM TRUE MEAN COMPARED WITH PROBABILITIES GIVEN BY NORMAL AND TYPE I CURVES

Data: Prices for Drug #24, Independent Stores, Manhattan, February, 1934

Limits (True Mean = 23.4¢)	Probability that the Mean of a Sample of Five Will Fall within Specified Limits as Given by		
	Histogram (Exact Probability)	Type I Curve	Normal Curve
Above 24.0	.3108	.3164	.2728
Below 22.8	.2447	.2571	.2539
Outside 22.8–24.0	.5555	.5735	.5268
Above 24.4	.1701	.1611	.1538
Below 22.4	.1528	.1584	.1387
Outside 22.4–24.4	.3230	.3195	.2925
Above 24.8	.0464	.0497	.0775
Below 22.0	.0831	.0904	.0641
Outside 22.0–24.8	.1295	.1402	.1416
Above 25.6	0	0	.0117
Below 21.6	.0417	.0477	.0233
Outside 21.6–25.6	.0417	.0477	.0350
Above 26.4	0	0	.0010
Below 20.4	.0034	.0045	.0008
Outside 20.4–26.4	.0034	.0045	.0018

These three non-normal types can be compared by the mean deviations from normality of the distributions of the means of samples of five. They are for the J-shaped, N-shaped, and skew normal: 0.011, 0.008, and 0.002 respectively.

The error involved in assuming normality of distribution of means of samples drawn from non-normal universes can be shown more clearly by comparing the fit of a normal curve to the distribution of means with the fit of a Pearsonian curve. This comparison was made and the results indicate that a Pearson Type I corresponds rather well

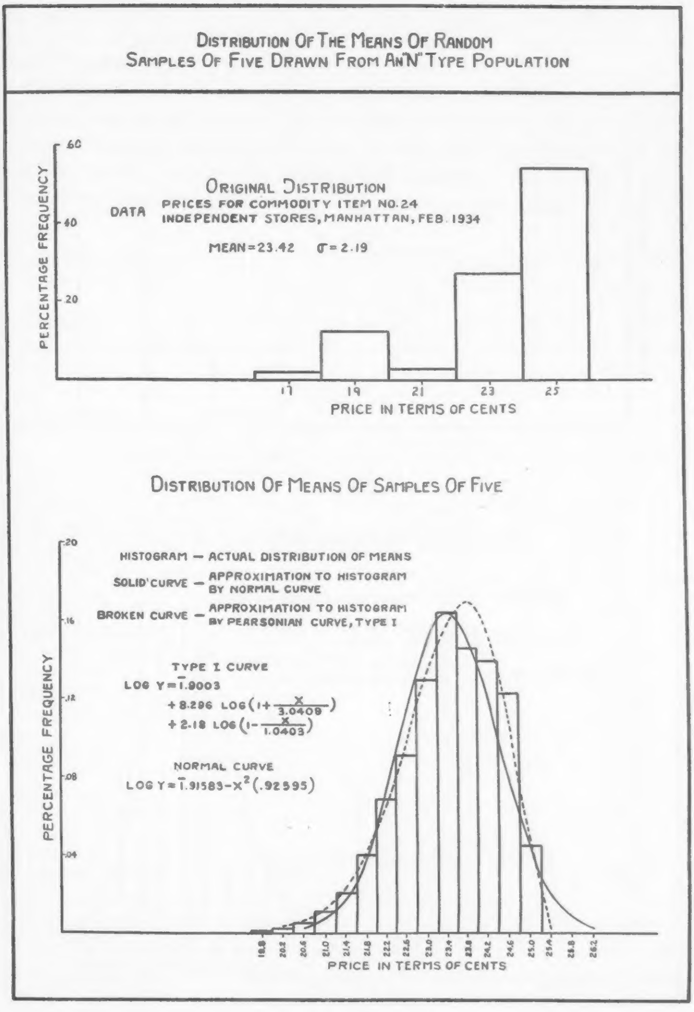


FIGURE 2.

with the distribution of means (cf. Figures 2 and 3). As might be expected from the preceding experiments, the improvement in fit of a Pearsonian over that of a normal curve increases as the original uni-

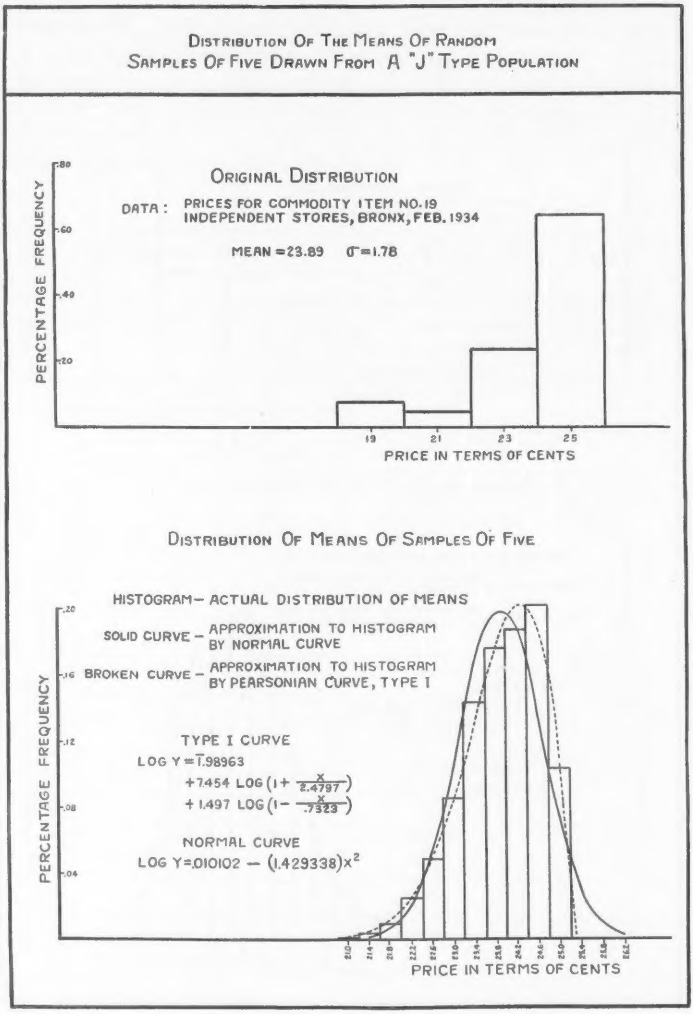


FIGURE 3.

verse changes from a near-normal to an N-shaped to a J-shaped. The percentage of error involved in using a normal curve to estimate the probability that a mean will lie outside of a given range increases as

the range is increased. The greatest improvement introduced by the use of the Type I curve is in the representation of asymmetry and in the tail points (Tables 5 and 6).

There remain to be found the increases in minimum sizes of samples made necessary by the very skew original distributions. It was shown earlier that errors in probability due to assumptions of normality are reduced rapidly as the size of sample is increased; the results permit conclusions concerning the size of the sample that should be drawn.

TABLE 6

EXACT PROBABILITIES OF DEVIATIONS OF MEANS OF RANDOM SAMPLES OF FIVE FROM TRUE MEAN COMPARED WITH PROBABILITIES GIVEN BY NORMAL AND TYPE I CURVES

Data: Prices for Drug #19, Independent Stores, Bronx, February, 1934

Limits (True Mean = 23.9¢)	Probability that the Mean of a Sample of Five Will Fall within Specified Limits as Given by		
	Histogram (Exact Probability)	Type I Curve	Normal Curve
Above 24.4	.3090	.2893	.2562
Below 23.2	.3251	.3081	.3513
Outside 23.2–24.4	.6341	.5974	.6075
Above 24.8	.1057	.1021	.1199
Below 22.8	.0941	.0939	.0787
Outside 22.8–24.8	.1998	.1959	.1986
Above 25.2	0	0	.0448
Below 22.4	.0435	.0437	.0261
Outside 22.4–25.2	.0435	.0437	.0709
Above 26.0	0	0	.0030
Below 22.0	.0172	.0179	.0069
Outside 22.0–26.0	.0172	.0179	.0100
Above 26.8	0	0	.0001
Below 20.8	.0005	.0004	...
Outside 20.8–26.8	.0005	.0004	.0001

If the universe is characterized by a near-normal distribution, a sample of five will be large enough to counteract this error to a large extent.[8] If the distribution is N-shaped, there is a decrease in the error as the sample size increases from three to four to five, but the error in a sample of five is still large,—much larger than in the case of the near-normal. Since the J-shaped is the worst offender, this has been examined more closely.

[8] This, of course, refers to the error in the standard error. The latter itself may remain large.

The prices of a certain drug in the neighborhood stores of Manhattan presented a J-shaped distribution. For this commodity, the expected frequencies for means of samples of five and seven were calculated by combinatorial method; and, as was true for the J-shaped universe discussed earlier, the means of samples of five follow a distribution that is far from normal. However, the distribution of means of samples of seven is rather close to normal.

To substantiate these theoretical results, one hundred random samples of five and one hundred of seven were drawn from this universe, and the distributions of the means of these two types were compared with their theoretical distributions. The correspondence was sufficiently close in both cases, there being an expectation of getting as bad or worse fit in thirty-four cases out of one hundred due to the fluctuations of random sampling for samples of five, and an expectation of forty-four out of one hundred for samples of seven. Then one hundred samples of ten were drawn from the same J-shaped universe; and the means of these samples were found to be distributed nearly symmetrically.[9] In fact, a normal curve follows the distribution much more closely than the theoretical histograms fit the other distributions of means; by a Chi-square test it was found that as bad or a worse fit might be expected in ninety-three cases out of one hundred. We are led to believe that a sample of ten is large enough to remove most of this kind of error from the standard error, and that a sample of seven is practically sufficient for this purpose. The standard error itself is reduced from 0.90 cents to 0.65 cents by increasing the number in the sample from five to ten.

The minimum adequate sample may be considered as the magnitude of a sample necessary to insure that, in at least ninety-five times out of one hundred, the sample mean will not differ by more than a specified amount from the true mean. Taking 10 per cent as the deviation permissible, the minimum adequate samples for twenty-six proprietary drugs were found.[10] With some exceptions, the outline of 1934 follows that of 1933, but in general the samples required are smaller in 1934 than in 1933. The approximate formula for the standard deviation of means was used, so sizes of samples listed might be in excess of those actually required. But also there should be an allowance for error in the sampling errors due to skew populations. Three of the items required samples of only three observations in 1934, but this must be qualified if it is known that their underlying distributions are very

[9] Table 7 presents the results of the experimental sampling.

[10] These are represented graphically in Table 8. It is to be noted that the minimum number is not strictly a minimum, but a minimum within the set of numbers 3, 5 8, 10, 12, and 15.

skew. In the absence of information to the contrary, one might expect them to be skew, in which case samples of five would be minimums; and there would be fair likelihood that they are J-shaped, thus requiring samples of seven. The same restriction applies to the items under a minimum requirement of five; for example, the distribution of the third item is known to be J-shaped. It is safer, in sampling for drug prices whose frequency distributions are unknown, to take samples of not less than six or seven.

TABLE 7

COMPARISON BETWEEN THEORETICAL FREQUENCY DISTRIBUTIONS OF MEANS OF SAMPLES DRAWN FROM J-SHAPED UNIVERSE AND DISTRIBUTIONS OBTAINED BY EXPERIMENTAL SAMPLING

Data: Original distribution of Drug #24, neighborhood independent stores, February, 1934; and three sets of one hundred samples drawn at random from this distribution

Distribution	Mean	Standard Deviation	Chi-Square	Probability
Original	23.13	2.34		
Theoretical for Samples of Five	23.18	1.01		
Observed for Samples of Five	23.04	.90	9.08	.34
Theoretical for Samples of Seven	23.17	.86		
Observed for Samples of Seven	23.09	.81	10.05	.44
Observed for Samples of Ten	22.99	.65		
Normal Fitted to Observed for Ten	22.99	.65	4.37	.93

IV. STANDARD ERRORS OF RETAIL PRICE INDEXES

It was suggested in the previous section that the average price of a commodity might be made up of a sum of prices weighted according to the quantities sold at those prices, for example, the average chain-store price weighted by chain-store sales plus the average neighborhood-store price weighted by neighborhood sales, and so forth. Therefore, a retail price index would involve a double weighting process: first, weighting the prices of a given commodity by volume of sales at those prices; and second, weighting the different commodities according to their importance in consumption. This line of analysis will be followed

in a later study, but the present section will be devoted more to a discussion of errors in the prices themselves than to errors in their weights.

The aspect of the index-number problem that is emphasized here is the effect of very skew distributions of component prices upon the sampling errors of the index number. It is well known that the vari-

TABLE 8

SIZE OF SAMPLE NECESSARY TO INSURE THAT IN NINETY-FIVE CASES OUT OF ONE HUNDRED THE MEAN OF THE SAMPLE WILL NOT DIFFER FROM THE TRUE MEAN BY MORE THAN TEN PER CENT

Data: Prices from All Stores, Manhattan, February, 1933 and 1934

Branded Drug Number	1933						1934					
	3	5	8	10	12	15	3	5	8	10	12	15
19	x						x					
10		x						x				
24		x						x				
25		x						x				
2			x						x			
6			x					x				
7			x						x			
8			x						x			
9			x					x				
11			x				x					
14			x				x					
15			x						x			
1				x							x	
12				x						x		
20				x					x			
16					x				x			
21					x				x			
22					x				x			
5						x			x			
17						x						x
18						x			x			
26						x				x		

ance of an aggregate is equal to the sum of the variances of the components, assuming independence; and the higher moments may be obtained in terms of the moments of the original distribution by the same sort of proof. But, since the probable error of each mean price which makes up the aggregate is in error because of a skew population, there is also an error in the probable error of the aggregate. This effect, it may be assumed, will be greatly neutralized by the following considerations:

1. The distributions of means obtained by sampling the individual commodity prices are more nearly normal than the original distributions; and

2. By the same reasoning, the distribution of the average of the various commodity mean prices is likely to be more nearly normal than the distributions of the mean prices themselves.

There is a third factor the importance of which it is difficult to estimate: the offsetting effects of some positively and some negatively skewed distributions.

To illustrate these tendencies, the theoretical distribution for an index consisting of the sum of means of random samples of five from two J-shaped distributions (drugs #19 and #24) was calculated by the combinatorial method. In order to have a basis for comparison, two distributions already examined were chosen. The normality of the index is compared with the normality of the individual samples composing the index by finding the sum of the deviations from the normal curve without regard to sign. Thus, the mean deviation of the index is 0.00363 as compared with the mean deviations of 0.00826 and 0.01122 for the individual samples.

To determine the effect of the third factor, which was not present in the preceding illustration (since both distributions were skewed in the same direction), the distribution of prices for drug #19 was reversed. The mean deviation from normality for this index of a J-shaped universe and a reverse J-shaped one is less than one-tenth of the mean deviation for the more normal of the component distributions; i.e., 0.000776. Moreover, it is about one-fifth of the mean deviation of the index without reversal.

The factor of the population distributions may be expected to have a great influence on an index made up of just two commodities, but, as the number of commodities included increases, the importance of these distributions decreases rapidly. Judging from the study of distribution of means of samples (which is an analogous case) an index of ten commodities should be large enough to eliminate this factor. In addition, the presence of reversed skewness for some commodities will help to eliminate the error in the index standard error. Increasing the sizes of the price samples accomplishes the same result. These conclusions were verified by a number of empirical tests.

JOHN H. COVER

University of Chicago
Chicago, Illinois

SOME A POSTERIORI PROBABILITIES IN STOCK MARKET ACTION

By ALFRED COWLES 3RD AND HERBERT E. JONES

IN 1933, one of the authors[1] published an analysis of the results secured by 45 representative financial agencies in forecasting the prices of common stocks. This study, embracing the period of $4\frac{1}{2}$ years from January, 1928 to June, 1932, and including several thousand individual forecasts, showed that these were unsuccessful more often than successful and that, indeed, better results in the aggregate would probably have been secured by investors through following purely random investment programs. This result naturally suggested the question: Is stock price action random in nature or, if not, to what extent is it possible statistically to define the nature of its structure? Herewith is presented at least a partial answer to this question on the basis of evidence adduced from internal elements in the stock price series themselves.

Among the different viewpoints from which this problem may be approached are those which consider (1) the element of inertia, and (2) harmonic analysis for the purpose of disclosing evidence as to regular periodicity.

With respect to the latter, Professor Harold T. Davis[2] has presented evidence of periodicity in stock prices by means of Schuster's periodogram analysis, and has attempted tests of the significance of these periods by R. A. Fisher's technique. By this method, however, only *average* periods can be found and even then their significance can not be accurately determined because of uncertainty as to the independence of observations. Variable periodicity can be taken into account by a number of techniques, such as Frisch's changing harmonics,[3] or by using varying moving averages,[4] but there seems to be no way to determine the probabilities of significance in connection with any of these methods. For this reason in the following analysis a method more susceptible of interpretation in terms of probability has been used.

Carl Snyder,[5] in referring to measurements of the industrial growth

[1] Alfred Cowles 3rd, "Can Stock Market Forecasters Forecast?", ECONOMETRICA, Vol. 1, 1933, pp. 309–324.

[2] In a paper presented before the Econometric Society in St. Louis, January, 1936. See abstract in ECONOMETRICA, Vol. 4, 1936, pp. 189–190.

[3] Ragnar Frisch, "Changing Harmonics and Other General Types of Components in Empirical Series," *Skandinavisk Aktuarietidskrift*, 1928, pp. 220–236.

[4] Gerhard Tintner, *Prices In the Trade Cycle*, Vienna, 1935, p. 23.

[5] "Concepts of Momentum and Inertia," presented at the Atlantic City meeting of the Econometric Society, December, 1932, and published in *Stabilization of Employment*, Charles F. Roos, editor, 1933, p. 77.

of the United States since 1830, says, "The picture that these measures give is that of an amazingly even rate of growth not merely from generation to generation but actually of *each separate decennium* throughout the last century. As if there was at work a kind of momentum or inertia that sweeps on in spite of all obstacles."

The term "inertia," as employed by Mr. Snyder, may be said to be macroscopic, whereas its use in the present study is microscopic. He was concerned with the trend over a period of 100 years or more and concluded that all deviations from that trend were nothing more than inconsequential "jiggles." The present analysis, on the other hand, is concerned primarily with evidence as to inertia in movements of a few hours, days, weeks, months, or years, which Mr. Snyder, with his longer-range viewpoint, would designate as mere "jiggles."

Evidence of inertia may be disclosed in the following manner. In a penny-tossing series there is a probability of one-half that tails will follow heads and vice versa. If the stock market rises for 1 hour, day, week, month, or year, is there a probability of one-half that it will decline in the succeeding comparable unit of time? In an attempt to answer this question sequences and reversals were counted, a sequence occurring when a rise follows a rise, or a decline a decline, and a reversal occurring when a decline follows a rise, or a rise a decline.

A study of the ratio of sequences to reversals will disclose structure as defined above, if it exists within the series, and the significance of this structure can be defined by ordinary statistical methods. For instance, the probability can be determined that any ratio occurred by chance from a random population.[6] Also, the consistency of these ratios can be investigated and from their frequency distributions one can determine the probabilities of success in forecasting a rise or decline in stock prices. Samples, of adequate length where available, were ex-

[6] L. Besson, "On the Comparison of Meteorological Data with Results of Chance," translated and abridged by E. W. Woolard, *Monthly Weather Review*, Vol. 48, 1920, pp. 89–94, pointed out that in a *random series*, the ratio of sequences to reversals will be 0.5, that is, there will be twice as many reversals as sequences. It should be noted that, in the present analysis, the data employed are not *random*, but rather *cumulated random series*, that is series in which the first differences, rather than the actual observations, are random. In a truly random series the auto-correlation drops to zero when the series is lagged against itself by so much as one observation. In a cumulated random series this is not the case. For example, in the stock price series under consideration, even when the first differences have been rearranged in a random manner, an auto-correlation with a lag of one observation will yield a very high coefficient. This correlation coefficient will be a function of the length of the series and is approximately equal to $1-\log n/(n-1)$. In such a series, with the first differences random, the ratio of sequences to reversals will be, if n is large, 1.0 instead of 0.5 as observed in the case of a series in which the observations themselves are random.

amined, the intervals between observations being successively 20 minutes, 1 hour, 1 day, 1, 2, and 3 weeks, 1, 2, 3, $\cdots$, 11 months, and 1, 2, 3, $\cdots$, 10 years. The results of this investigation are presented in Table 1 and are shown graphically in Figure 1.

It was found that, for every series with intervals between observations of from 20 minutes up to and including 3 years, the sequences outnumbered the reversals. For example, in the case of the monthly series

TABLE 1

RATIO OF SEQUENCES TO REVERSALS IN STOCK PRICE INDEXES

UNIT	INDEX	PERIOD	NUMBER OF OBSERVATIONS	RATIO OF SEQUENCES TO REVERSALS	PROBABILITY OF OCCURANCE BY CHANCE	UNIT	INDEX	PERIOD	NUMBER OF OBSERVATIONS	RATIO OF SEQUENCES TO REVERSALS	PROBABILITY OF OCCURANCE BY CHANCE
20 MINUTES	HARRIS UPHAM	1935-1936	2800	1.44	<.000001	8 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	156	1.48	.01640
1 HOUR	DOW JONES HOURLY AVGS	1933-1934	800	1.29	.00040	9 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	138	1.57	.01016
1 DAY	DOW JONES HOURLY AVGS.	1931-1935	1200	1.18	.00094	10 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	124	1.49	.03000
1 WEEK	STANDARD STATISTICS	1918-1935	938	1.24	.00386	11 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	113	1.27	.21870
2 WEEKS	DOW JONES	1897-1935	976	1.02	.80258	1 YEAR	INDEX OF RAILROAD STOCK PRICES	1835-1935	100	1.17	.42952
3 WEEKS	DOW JONES	1897-1935	652	1.08	.30772	2 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1935	50	1.63	.08726
1 MONTH	INDEX OF RAILROAD STOCK PRICES	1835-1935	1200	1.66	<.000001	3 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1934	33	1.46	.28914
2 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	600	1.50	<.000001	4 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1935	25	0.85	.68180
3 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	400	1.29	.01242	5 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1935	20	1.00	1.00000
4 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	300	1.18	.16452	6 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1931	16	0.67	.44130
5 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	249	1.52	.00120	7 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1933	14	0.71	.56192
6 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	208	1.40	.01778	8 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1931	12	0.22	.03486
7 MONTHS	INDEX OF RAILROAD STOCK PRICES	1835-1935	178	1.38	.03486	10 YEARS	INDEX OF RAILROAD STOCK PRICES	1835-1935	10	0.60	.74140

The Index of Railroad Stock Prices is composed of several series which were assembled and put together by the Cleveland Trust Company. Canal stock prices were used for the period 1831 through 1833. From 1834 through 1879 the index is based on three Harvard series. The one from 1834 through 1852 includes eight stocks, and that from 1853 through 1865 includes 18. The data are from *The Review of Economic Statistics* for August, 1928. The index from 1866 through 1879 includes 10 stocks, and the data are from *The Review of Economic Statistics* for 1919. The index from 1880 through 1896 includes 10 stocks, and from 1897 to date it includes 15. These two latter indexes were compiled by the Cleveland Trust Company. All the earlier indexes were adjusted to form a continuous series terminating with the final index of 15 stocks.

from 1835 to 1935, a total of 1200 observations, there were 748 sequences and 450 reversals. That is, the probability appeared to be .625 that, if the market had risen in any given month, it would rise in the succeeding month, or, if it had fallen, that it would continue to decline for another month. The standard deviation[7] for such a long series

[7] In a random penny-tossing series the probability of a sequence or reversal is

constructed by random penny tossing would be 17.3; therefore the deviation of 149 from the expected value 599 is in excess of eight times the standard deviation. The probability of obtaining such a result in a penny-tossing series is infinitesimal. If the unit of time be increased to 6 months, we find that there are 120 sequences to 86 reversals or what appears to be a .583 probability that a sequence will occur in any successive pair of periods. The probability in this case is

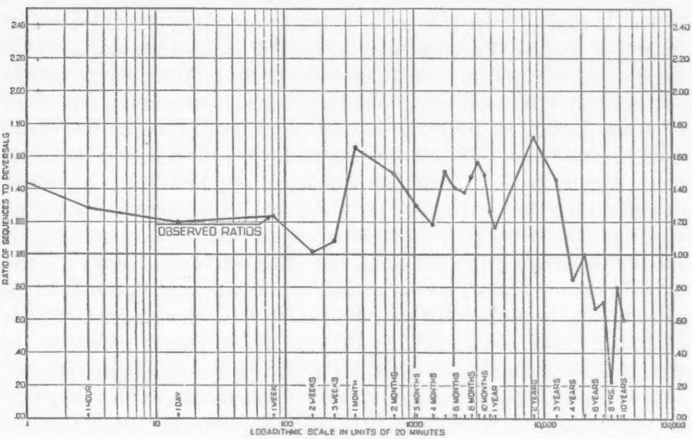


FIGURE 1.—Ratio of sequences to reversals in direction of stock price indexes for various time intervals.

.01778 that such a ratio of sequences to reversals might occur in a random series such as that of penny tossing previously referred to.

For annual series, the probability for chance occurrence, on the same basis as before, is .42952. In fact, the probabilities are inconclusive for all series using units of over 6 months, although this may be due to the limitations of the data. There seems to be a good chance that with more data it might be possible to demonstrate the existence of struc-

$\frac{1}{2}$ and the standard deviation is $\sqrt{npq}$ or $\sqrt{n}/2$ where n is the number of observations and p and q are respectively the probabilities of success and failure. Sequences or reversals are determined from the first differences of the original price series. The total number of sequences and reversals is one less than the number of first differences. The standard deviation of the number of sequences, therefore, will be $\sqrt{n-2}/2$ where n is the number of observations of the original price series. The actual deviation will be $S-(n-2)/2$ where S is the observed number of sequences and the ratio $[S-(n-2)/2]/[(n-2)^{\frac{1}{2}}/2]$ is the observed deviation in terms of the standard deviation. The probability of as large a deviation occurring by chance can be found in the ordinary table of the normal probability function.

ture for all units of time up to and including 3 years. It is difficult, however, to explain the very small excess of sequences over reversals for series using intervals of 2 weeks and 3 weeks in view of the fact that for the slightly shorter unit of 1 week we have significant indication of structure and also for series using intervals of 1 month. In fact, as will be shown later, the series represented by units of 1 month proves to be the most significant from a practical point of view.

The above analysis was based upon a study of the stock market as a whole. It may now be of interest for us to examine the evidence with regard to industrial groups such as motors, oils, steels, and so forth. To this end the indexes of common stock prices of 61 industrial groups, prepared by the Standard Statistics Company, were analyzed. Their monthly deviations from the median were noted, the median being used because more groups are normally below, than above, the arithmetic means, in view of the fact that extremely large gains are larger in percentage than extremely large declines. Sequences and reversals were counted for the purpose of determining whether there was a tendency for such groups to persist in exceeding, or falling below, the median. In other words, if the oil stocks, as a group, were in the 30 of the 61 industrial groups which advanced more than the median group in January, what is the probability that the oils will also be found in the strongest 30 groups in February? For the 16 months from January, 1934 to April, 1935, when the general market movement was approximately horizontal, in 917 observations there were 570 sequences and 345 reversals. That is, if the oil stocks were among the strongest 30 groups in January the probabilities would appear to be .623 that they would also be found among the strongest 30 groups in February. The application of the theory of probability to interpret the significance of this result is hampered, as in many other analyses of economic time series, by uncertainty as to what is the number of independent observations in the sample. The action of the oils, for example, may be correlated with that of the motors or some other group, so that when one is stronger than the median, the other also tends to be stronger. Modern statistical technique appears to offer no ready solution for this problem of the independence of observations, and we must, therefore, content ourselves with obtaining unusually favorable probabilities.

In the period, May, 1935 to February, 1936, a rising market of 551 observations, there were 379 sequences and 170 reversals, indicating an apparent probability of .690 in favor of a sequence. Here, however, another factor has intruded itself. The stocks of certain industries can be shown to be more cyclical in nature than those of other industries. For instance, in severe depressions the building of new houses is almost completely stopped, and yet people under such conditions go on eating

almost as much food as in periods of prosperity. These tendencies are reflected by wider fluctuations in the earnings of producers of building materials than in the case of purveyors of food, and also by wider cyclical fluctuations in the stocks of the former corporations than of the latter. This tendency results, in the case of a rising market, in the building-stock prices being persistently stronger than the average, and the food stocks persistently weaker. A count of sequences and reversals under such conditions measures, therefore, to some extent, the differences in cyclical behavior among the various groups rather than the tendency toward inertia which is measured in the case of a period where the market as a whole is moving horizontally. Since there are not many periods of great length in which the market as a whole has moved horizontally, we are limited in the data available for this particular analysis. Considering only the data for the horizontal 16 months from January, 1934 to April, 1935, the excess of sequences over reversals is 7.5 times the standard error for a random series. Even assuming that half of the observations are not independent, there is still the exceedingly small probability of .00168 of occurrence in a random series.

The action of individual stocks was also investigated. Instead of the oils as a group, it was considered, for example, whether the Standard Oil Company of New Jersey, if it were stronger than the median of all stocks in January, would more likely than not be stronger in February. Taking 190 representative stocks for the years 1934 and 1935 and the first three months of 1936, and using the same technique as that employed in the case of the industrial groups, a total of 4659 observations, there was found, when the market moved horizontally, from January, 1934 to May, 1935, a ratio of sequences to reversals of 1.07 to 1 which is 9 times the standard error. In the last 12 months, a rising period, the ratio of sequences to reversals was 1.29 to 1. It follows from the previous discussion of cyclical behavior that the difference between the two periods is to be expected and that the evidence of the second period must be brought in question.

Taking 1 year as the unit of measurement for the period from 1920 to 1935, the tendency is very pronounced for stocks which have exceeded the median in one year to exceed it also in the year following. In 1837 observations were 1200 sequences and 635 reversals. The excess of sequences was about 13 times the standard error for a random series constructed on the basis of equal probabilities. During the period under consideration the market as a whole manifested 8 sequences and 7 reversals. The evidence therefore seems to indicate that, when 1 year is the unit of time, individual stocks will persist in doing better, or worse, than the median of all stocks.

This evidence of structure in stock prices suggests alluring possibilities in the way of forecasting. In fact, many professional speculators, including in particular exponents of the so-called "Dow Theory" widely publicized by popular financial journals, have adopted systems based in the main on the principle that it is advantageous to swim with the tide. The practicability of such forecasts, however, will depend,

TABLE 2

ABSOLUTE PERCENTAGE CHANGES IN STOCK PRICE INDEXES

UNIT	INDEX	PERIOD	NUMBER OF OBSERVATIONS	AVERAGE ABSOLUTE CHANGE IN PERCENT	STANDARD DEVIATION OF AVERAGE
20 MINUTES	HARRIS-UPHAM	JULY-9-1936 JULY-17-1936	111	0.12	.01
1 HOUR	DOW JONES HOURLY AVGS.	SEPT-12-1935 OCT-6-1935	102	0.32	.03
2 HOURS	DOW JONES HOURLY AVGS.	AUG-1-1935 OCT-6-1935	103	0.47	.04
1 DAY	DOW JONES	AUG-27-1934 DEC-31-1934	102	0.73	.07
1 WEEK	DOW JONES	JAN-6-1913 DEC-31-1934	1128	2.56	.21
1 MONTH	DOW JONES	JAN-1-1897 DEC-31-1934	451	3.70	.46
2 MONTHS	DOW JONES	APRIL-1-1918 DEC-1-1934	100	5.02	.91
3 MONTHS	INDEX OF RAILROAD STOCK PRICES	JAN 1835 DEC 1934	400	8.92	.79
4 MONTHS	DOW JONES	DEC 1900 DEC 1934	100	10.79	.99
5 MONTHS	INDEX OF RAILROAD STOCK PRICES	JAN 1893 SEPT 1934	100	8.62	.82
6 MONTHS	INDEX OF RAILROAD STOCK PRICES	DEC 1884 DEC 1934	100	10.04	1.20
7 MONTHS	INDEX OF RAILROAD STOCK PRICES	JUNE 1876 OCT 1934	100	11.81	1.30
8 MONTHS	INDEX OF RAILROAD STOCK PRICES	APRIL 1868 DEC 1934	100	11.30	1.13

UNIT	INDEX	PERIOD	NUMBER OF OBSERVATIONS	AVERAGE ABSOLUTE CHANGE IN PERCENT	STANDARD DEVIATION OF AVERAGE
9 MONTHS	INDEX OF RAILROAD STOCK PRICES	JUNE 1859 JUNE 1934	100	12.73	1.29
10 MONTHS	INDEX OF RAILROAD STOCK PRICES	JAN 1851 APRIL 1934	100	13.00	1.31
11 MONTHS	INDEX OF RAILROAD STOCK PRICES	DEC 1842 JULY 1934	100	13.99	1.25
1 YEAR	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1934	103	14.70	1.43
2 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1933	51	22.58	2.78
3 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1933	34	28.03	4.81
4 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1932	25	30.59	4.77
5 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1931	20	33.95	5.71
6 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1933	17	38.59	8.90
7 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1829	14	33.54	9.62
8 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1927	12	32.38	9.15
9 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1930	11	45.98	13.64
10 YEARS	INDEX OF RAILROAD STOCK PRICES	JAN 1831 JAN 1931	10	51.64	10.90

not only on the ratio of sequences to reversals, but also on the brokerage costs and the average change in stock prices during the unit of time selected. The brokerage costs, of course, are known. To determine the average percentage change in the stock prices for various units of time an extensive study was made. The difference between the index at the beginning of one unit and the beginning of the next, given in percentage of the former value, was computed. The results are shown in Table 2. These averages are absolute values, that is, the directions of the moves were not considered. The values, plotted on a time scale, are shown in Figure 2. The increase is very regular, approximating a smooth exponential curve as shown in the diagram.

With these data it is possible to compute the net gain which would have resulted from an application of this type of forecasting to the stock market averages.

Let

$I(t)$ = Expected annual net profit, in per cent,
$R(t)$ = Ratio of sequences to reversals for time interval t,
$C(t)$ = Average change per time interval t in per cent,
$Y(t)$ = Number of time intervals, t, in one year,
B = Brokerage cost for one complete trade, in per cent.

In the long run there will be an average of $R(t)$ sequences for each reversal, that is, a speculator will, on the average, be on the right side of the market for $R(t)$ units of time for each unit that he is on the

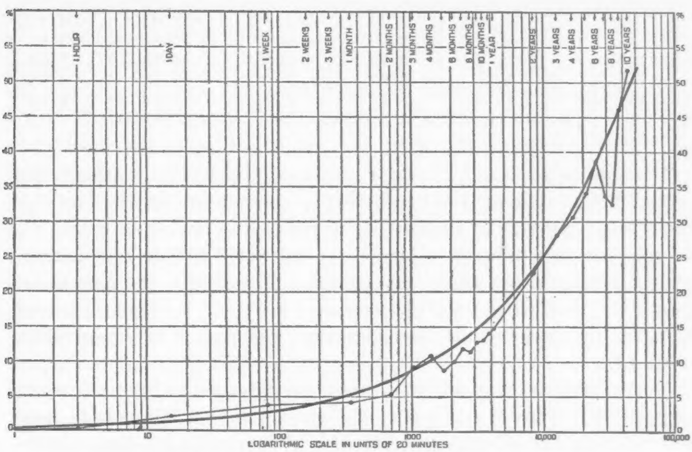


FIGURE 2.—Absolute percentage changes in stock price indexes.

wrong side, if his market position is changed only *after* the occurrence of each reversal. The average net time in the right direction between changes of position will then be $[R(t)-1]$ time units. The average move per unit of time is $C(t)$; therefore, the *gross* gain per position will then be $[R(t)-1]C(t)$. The *net* gain per position will be

$$[R(t) - 1]\cdot C(t) - B.$$

Since one will be in the market in the right direction $R(t)$ units of time and in the wrong direction 1 unit, the total time per position will be $[R(t)+1]$. The number of positions taken per year will be $Y(t)/[R(t)+1]$. The net annual gain, in per cent, will then be given by

$$I(t) = 100\left[\left\{[R(t) - 1]C(t) - B\right\}\frac{1}{100} + 1\right]^{Y(t)/[R(t)+1]} - 100.$$

When the values of $R(t)$ and $C(t)$ are substituted in this equation values of the expected annual net gain for various units of time and brokerage costs are obtained as shown in Table 3.

TABLE 3

Time Unit	Ratio of Sequences to Reversals $R(t)$	Average Percentage Changes $C(t)$	Expected Annual Net Profit		
			Brokerage Costs of:		
			1%	$1\frac{1}{2}$%	2%
1 day	1.18	0.73%	−67.4%	−83.0%	−91.1%
1 week	1.24	2.56	− 8.55	−18.6	−27.6
1 month	1.66	3.70	6.66	4.25	2.00
2 months	1.50	5.02	3.66	2.44	1.23
3 months	1.29	8.92	2.79	1.91	1.03

As might be expected, the daily and weekly units are too short. The probability of success is not sufficient to compensate for the fact that the changes per unit of time are small relative to brokerage costs. The average net gain *per trade* is largest for units of 2 months but, because with this unit so few trades are completed in a year, the annual net gain is less than when units of one month are used. It appears, indeed, that, for the period under consideration, one month is the optimum unit of time.

At this point a study was undertaken to determine whether it would be possible to select a group of stocks which would continue to be more volatile than the average and whether, if this could be done, the results of speculation, employing such a list, would be more successful than where average stocks were used. An investigation was made covering the period from 1900 to 1919. The number of stocks examined was 44 at the beginning of the period, increasing to 81 by 1918. The 10 per cent manifesting the greatest movement for 1 year, in the direction of the market, was chosen as the volatile group. It was found that such a group was about 1.7 times as volatile as the market for 6 months following its selection. The list of stocks making up the volatile group was adjusted at the end of each 6-month period.

The list of rapidly moving stocks, selected as described above, was then subjected to a comparison with the results secured employing average stocks for the period under consideration. The average stocks would have yielded an average annual net gain of 8.9 per cent while the volatile stocks would have shown an average net gain of but 5.4 per cent. Though the volatile group manifested an average move considerably more than that of the market, this was more than offset by the

fact that its ratio of sequences to reversals was less favorable. The substitution of volatile for average stocks was therefore abandoned.

In the case of the stock market averages a study was undertaken to determine the degree of consistency in the data by considering distributions of the variables over finite periods of time. Figure 3 represents the frequency distributions of the ratios of sequences to reversals for units of 1 day, 1 week, 1 month, and 3 months. Periods of 12 and 24 units of time were considered in all cases except the last, where insufficient data made it necessary to consider periods of only 4 units. In the upper part of each diagram are the actual histograms, or frequency distributions. A better way to illustrate the probability of variations is by summation, or cumulative frequency distribution. The histograms, therefore, were smoothed, as shown by the solid continuous curve, and the cumulated curves computed from the smoothed distributions. The dotted curves represent the distributions that would be expected from series for which the probability of a sequence or reversal is $\frac{1}{2}$. They were computed from the formula

$$P(S) = \frac{{}_{n-2}C_S}{2^{n-2}},$$

where $P(S)$ is the probability of obtaining S sequences in a sample of n observations.[8]

For units of 1 day the periods of 12 units have a distribution very nearly random with a slight tendency towards a skewness to the right. This skewness is more pronounced in the diagram for periods of 24 units.

The cumulative frequency curves have a scale on the left giving the probabilities of a smaller ratio, and a scale on the right, the probabilities of a larger ratio. From the figure it will be seen that the probability of obtaining a ratio of sequences to reversals less than 1 is about .43. Or again, the probability of obtaining a ratio larger than 2 is about .20 in the case of periods of 12 days and about .09 for periods of 24 days. From such a diagram it is possible to determine the limiting ratio for any probability. For instance, the limits of a probability band of .50 can be determined by finding the ratios for which the probabilities are .75 and .25, respectively, of obtaining a larger ratio. In this case, for

[8] In n observations there are $(n-1)$ first differences. In this set of $(n-1)$ first differences there can be S sequences and $(n-2)-S$ reversals. The number of different orders in which $(n-2)$ things can be arranged in two sets, S and $(n-2)-S$, is $(n-2)!/S!(n-2-S)! = {}_{n-2}C_S$. But a sequence can occur when either a rise follows a rise or a decline follows a decline. The total number of samples containing S sequences, therefore, will be $2 \cdot {}_{n-2}C_S$. Since there are 2^{n-1} possible samples, the probability of obtaining S sequences will be given by the above equation.

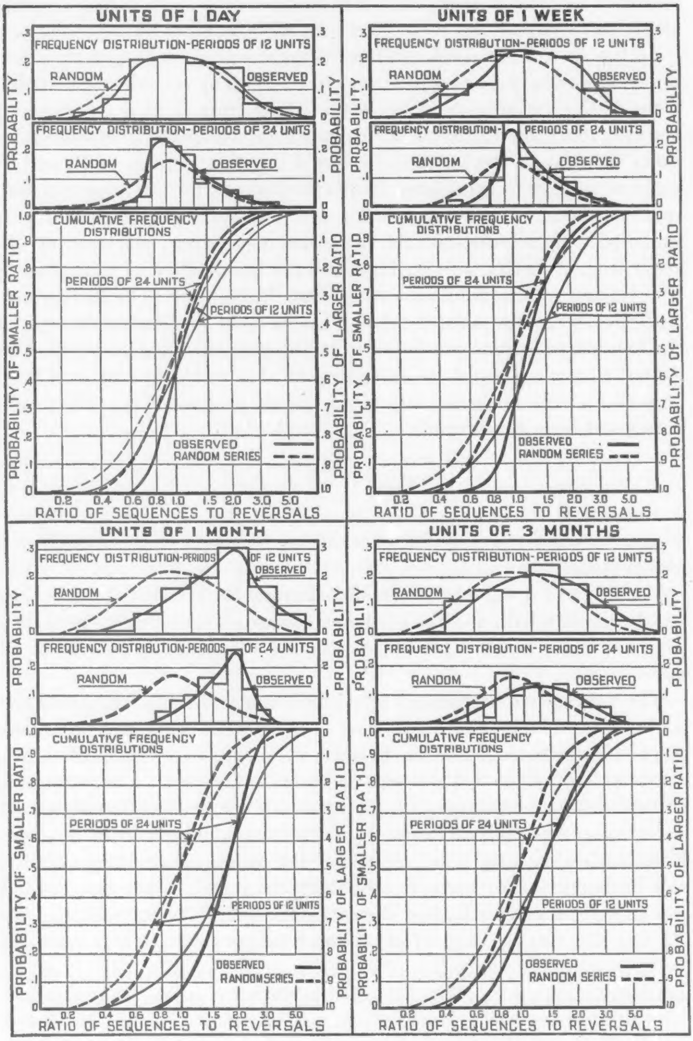


FIGURE 3.—Frequency distributions of ratio of sequences to reversals in direction of stock price indexes.

periods of 24 units, 0.9 is the lower limiting ratio and 1.4 the upper ratio. The median ratio is 1.1. Therefore, the ratio of sequences to reversals, for periods of 24 days, is 1.1; and 50 per cent of the time the ratios will lie between .9 and 1.4.

The distribution for units of 1 week shows a greater tendency towards structure than in the case of daily units. In this case the cumulative curve indicates that the probabilities are about 2 to 1 in favor of a ratio greater than unity. The tendency towards structure is even more

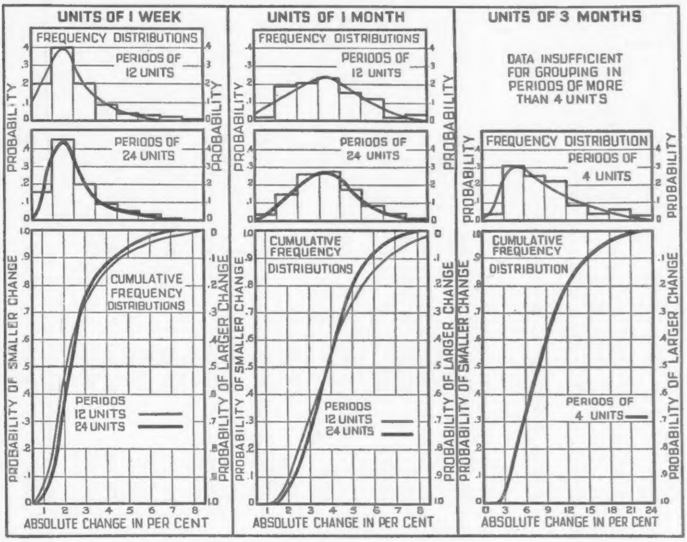


FIGURE 4.—Frequency distributions of absolute percentage changes in stock price indexes.

pronounced for units of 1 month where, for periods of 12 months, the chances are about 4 to 1 in favor of a ratio greater than unity. The cumulative curves are relatively steep, also, which indicates a high concentration around the average. This is especially true for the periods of 24 months. For units of 3 months the distributions are not nearly so favorable as in the previous case. They are flatter, indicating large variations, and the average ratio is much smaller.

Figure 4 shows the results of the same type of analysis applied to the data for the percentage changes. In this case only three units of time were investigated as the daily units already had proved to be impracti-

Table 4

Computation of Distribution of $I(t) = 100\left[\left\{[R(t)-1]C(t)-B\right\}\frac{1}{100}+1\right]^{Y(t)/[R(t)+1]} - 100$

For Units of 1 Month, i.e., $Y(t) = 12$, and Brokerage $(B) = 1\%$
Roman type figures within table give values of $I(t)$
Italic type figures within table give probabilities of $I(t)$

Ratio of Sequences to Reversals—$R(t)$	Average Percentage Changes—$C(t)$: 1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5	Probability of $R(t)$
.50	−13.17 *.0015*	−16.65 *.0042*	−20.00 *.0054*	−23.23 *.0043*	−26.34 *.0025*	−29.35 *.0010*	−32.25 *.0006*	−35.04 *.0004*	*.020*
.71	− 9.65 *.0060*	−11.50 *.0166*	−13.32 *.0216*	−15.10 *.0174*	−16.85 *.0100*	−18.57 *.0040*	−20.26 *.0024*	−21.92 *.0016*	*.080*
1.00	− 5.85 *.0109*	− 5.85 *.0302*	− 5.85 *.0392*	− 5.85 *.0315*	− 5.85 *.0181*	− 5.85 *.0072*	− 5.85 *.0044*	− 5.85 *.0029*	*.145*
1.40	− 1.98 *.0172*	0 *.0478*	2.02 *.0621*	4.07 *.0499*	6.15 *.0288*	8.26 *.0115*	10.41 *.0069*	12.59 *.0046*	*.230*
2.00	2.02 *.0029*	6.14 *.0634*	10.38 *.0824*	14.75 *.0662*	19.25 *.0381*	23.88 *.0152*	28.65 *.0092*	33.55 *.0061*	*.305*
3.00	6.12 *.0112*	12.49 *.0312*	19.10 *.0405*	25.97 *.0326*	33.10 *.0188*	40.49 *.0075*	48.15 *.0045*	56.09 *.0030*	*.150*
5.00	10.25 *.0038*	18.81 *.0104*	27.69 *.0135*	36.89 *.0108*	46.41 *.0062*	56.25 *.0025*	66.41 *.0015*	76.89 *.0010*	*.050*
Probability of $C(t)$	*.075*	*.208*	*.270*	*.217*	*.125*	*.050*	*.030*	*.020*	

cable. For units of 3 months the data were insufficient for grouping in periods of more than 4 units. For units of 1 week and 1 month the difference between the periods of 12 units and 24 units is so slight as to be without statistical significance.

The next step was to combine these distributions in order to determine the distribution of the expected net gain. To do this mathematically gives rise to hopeless complications unless certain simplifying assumptions are made which themselves cast doubt on the value of the solutions. Therefore, it was necessary to compute the distributions by empirical means.

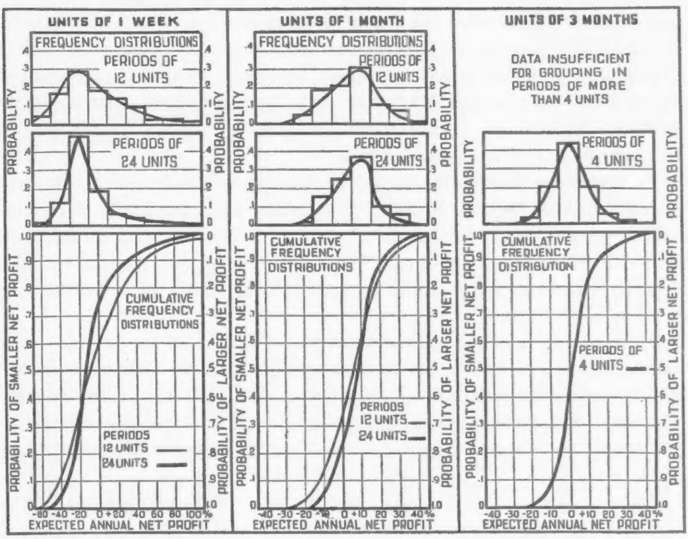


FIGURE 5.—Frequency distributions of expected annual net profits based on method of sequence probabilities. Brokerage computed at 1 per cent.

To determine these empirical distributions, values of $R(t)$ and $C(t)$ were taken which adequately covered their range of variation, and $I(t)$ was computed for all possible combinations of these values. The average brokerage charge was assumed to be 1 per cent per trade. To facilitate computation tables were employed of the type illustrated in Table 4.

The roman type figures in the body of the table give the values of $I(t)$ for the various values of $R(t)$ in the first column and $C(t)$ in the

first row. The probabilities for each value of $R(t)$ and $C(t)$ are shown in the last column and the bottom row respectively. The probabilities for the various values of $I(t)$ are the products of the probabilities of $R(t)$ and $C(t)$ and are shown by the figures in italics in the table.[9] The values of $I(t)$ have been classified and the sum of the probabilities for each value of $I(t)$ in any class taken as the probability of that class. The results of this analysis are presented in Figure 5.

Here again are shown the histograms and cumulative frequency curves representing units of 1 week, 1 month, and 3 months. For units of 1 week the cumulative curves indicate an average annual loss of about 10 per cent, with but 1 chance in 3 of obtaining a net profit over any period of 24 weeks. The same is true to a lesser extent when we use units of 3 months. In this case we see there is about an even chance of a loss or gain. For units of 1 month, however, an average net gain of about 7 per cent is indicated. But, even here, no great consistency is evident. In fact the cumulative curve indicates that the chance of loss for any one year is about 1 in 3.

Furthermore the results should be interpreted with caution in view of the fact that various units of time, other than 1 month, were considered and rejected. In all, 26 such units, ranging from 20 minutes up to 10 years, were examined. The series represented by units of one month, therefore, was selected by hindsight as the most favorable one in 26 trials.

This type of forecasting could not be employed by speculators with any assurance of consistent or large profits. On the other hand, the significant excess of sequences over reversals for all units from 20 minutes up to 6 months, with the exception of units of 2 weeks and 3 weeks mentioned previously, represents conclusive evidence of structure in stock prices.

ALFRED COWLES 3RD
HERBERT E. JONES

Cowles Commission for
Research in Economics
Colorado Springs, Colorado

[9] The values of $R(t)$ and $C(t)$ for the periods studied are randomly distributed and a correlation analysis between $R(t)$ and $C(t)$ gave a coefficient of -0.2 with a standard error of 0.1. This coefficient is on the borderline of significance but is very small. Therefore, the probabilities of $I(t)$ may be computed in this manner without great error.

THE DEMAND FOR BICYCLES IN THE NETHERLANDS[1]

By J. B. D. Derksen and A. Rombouts

Consecutively we shall discuss:

a. The demand for bicycles. Attention will be paid:
   1. To the structural development of the use of bicycles in our country;
   2. To the influence of the trade cycle upon demand.

b. The home production and the imports of bicycles into the Netherlands. The chief factors which determine their interrelation are investigated.

c. The exports of bicycles from the Netherlands.

I. THE DEMAND FOR BICYCLES

In considering the demand for an article such as the bicycle, we have to distinguish between the *demand for replacement* of bicycles put out of use because of defects, etc. and *the demand for new purchases* (i.e., by new consumers).

Table 1

Development of the Use of Bicycles in the Netherlands

Year	Number of bicycles (thousands)	Population (January 1) (thousands)	Bicycles per 1000 inhabitants	Year	Number of bicycles (thousands)	Population (January 1) (thousands)	Bicycles per 1000 inhabitants
1899	94	5,075	19	1917	859	6,583	131
1900	113	5,140	22	1918	872	6,725	129
1901	133	5,179	26	1919	861	6,799	127
1902	159	5,263	30				
1903	188	5,347	35				
1904	227	5,431	42	1925	1,811	7,315	248
1905	274	5,510	50	1926	2,223	7,416	300
1906	325	5,591	58	1927	2,266	7,526	301
1907	378	5,672	66	1928	2,325	7,626	305
1908	435	5,747	76	1929	2,485	7,731	321
1909	487	5,825	84	1930	2,627	7,832	335
1910	540	5,898	92	1931	2,755	7,953	346
1911	592	5,946	100	1932	2,834	8,062	352
1912	647	6,022	108	1933	2,893	8,183	354
1913	708	6,114	116	1934	3,020	8,290	364
1914	767	6,213	123	1935	3,206	8,392	382
1915	789	6,340	124	1936	3,363	8,475	397
1916	811	6,449	126				

[1] This paper contains the results of a study made in the Section for Research of the Business Cycle of the Netherlands Central Bureau of Statistics, The Hague, Holland. We may mention here the numerous investigations made by this Bureau by or under direction of Prof. Tinbergen (now in Geneva) and published in *De Nederlandsche Conjunctuur*. We are much indebted to Mr. J. M. Fleming, Geneva, for his kindness in correcting the English manuscript.

It should be emphasized that these categories concern the population as a whole. Therefore it is of no importance to our investigation if someone sells his bicycle to somebody else, who makes further use of it.

It follows from the definition that the new purchases in any year are equal to the increase of the total number of bicycles in use. For each year this number can be seen from the statistics on taxes, except for the period 1919–1924, when no tax was imposed on bicycles. The figures are shown in Table 1, which also gives the number of bicycles per 1000 inhabitants. Whereas this number was 248 per 1000 in 1925, it is now 397 per 1000 and there are no signs yet that this rapid increase will come to an end.

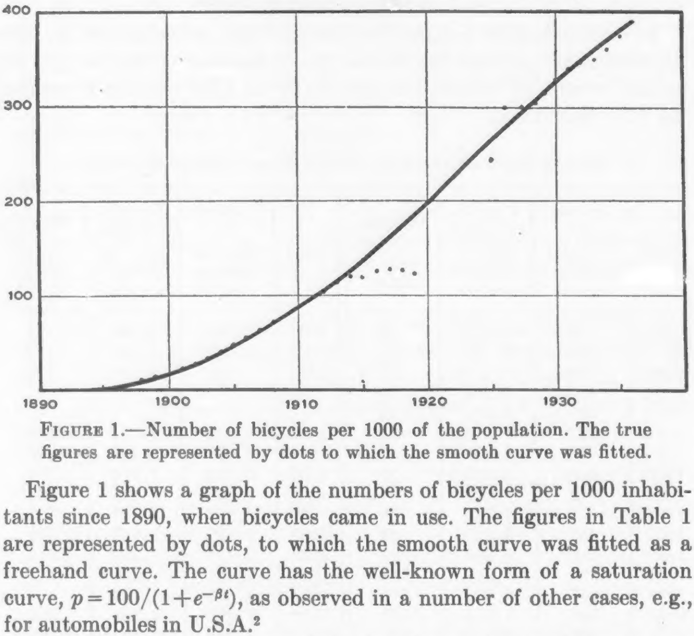


FIGURE 1.—Number of bicycles per 1000 of the population. The true figures are represented by dots to which the smooth curve was fitted.

Figure 1 shows a graph of the numbers of bicycles per 1000 inhabitants since 1890, when bicycles came in use. The figures in Table 1 are represented by dots, to which the smooth curve was fitted as a freehand curve. The curve has the well-known form of a saturation curve, $p=100/(1+e^{-\beta t})$, as observed in a number of other cases, e.g., for automobiles in U.S.A.[2]

From Figure 1 we may conclude that the market in Holland is not yet saturated.

One may naturally ask, whether this can be expected within a short time, which is practically the same as asking at what point saturation will be attained. A reliable result cannot be obtained as only too small a part of the saturation curve is known.

[2] Cf. *De Nederlandsche Conjunctuur*, Nov., 1936, p. 18.

We may guess at about 550 bicycles per 1000 inhabitants as the saturation point, an estimate which is supported by the data for Denmark, where the number of bicycles is estimated at 2 millions on a population of less than 4 millions.

The annual new purchases can be calculated from the increase of the total number of bicycles in use, which number may be derived from tax figures. This increase will be smaller than the sales. The difference between total sales and increase of the number of bicycles in use represents the number of bicycles for replacement. Those figures are found in Table 2.

TABLE 2

THE STRUCTURE OF THE DEMAND FOR BICYCLES

	1927	1928	1929	1930	1931	1932	1933	1934	1935
Total sales (in thousands)......	318	366	395	390	318	299	417	419	403
New purchases.......	59	160	142	128	79	59	127	186	157
Replacement purchases	259	206	253	262	239	240	290	233	246

From Table 2 we may conclude that the greater part of the sales is required for replacement. By the time the number of bicycles has ap-

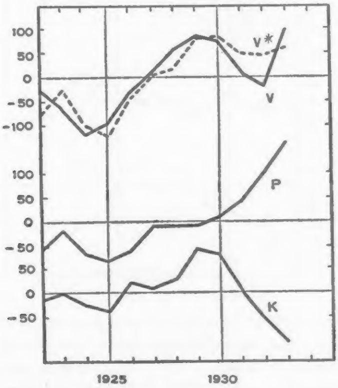


FIGURE 2.—V =total annual consumption.
V^* =the same, as calculated from formula (1).
P =influence of the price upon consumption.
K =influence of the purchasing power upon consumption.
All variables are represented as deviations from their average during 1922–1933.

proached the level of saturation, the new purchases will have begun gradually to decrease. So long as the population grows, a certain number of new purchases will be necessary.

From the figures given in Tables 1 and 2 we may calculate that the mean length of life of a bicycle amounts to 8 or 9 years.

After the above explanation of the structure of the demand for bicycles, we will now investigate how far the annual consumption of bicycles is influenced by the trade cycle. The result of the calculation was that the price of bicycles and the real income of the consumers are the most important factors influencing consumption. The following relation was derived from the data of the period 1922–1933:

$$V = 11.2K - 8.6P - 379. \tag{1}$$

In this formula V represents the total annual consumption of bicycles (in thousands), K is an index of the purchasing power of the consumers, and P the price of bicycles (in guilders). From Figure 2 can be seen that the above formula gives a satisfactory explanation of the course of consumption (the correlation coefficient turned out to be 0.88). In considering the course of the true and the calculated con-

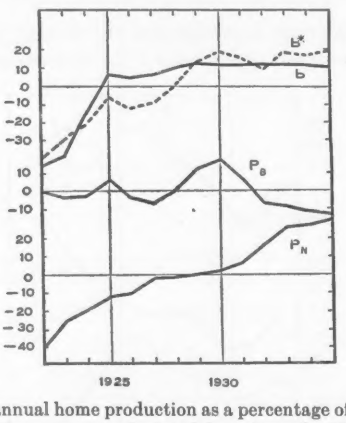


FIGURE 3.—b = annual home production as a percentage of total consumption.
b^* = the same, as calculated from formula (2).
P_N = influence of the price of Dutch bicycles.
P_B = influence of the price of foreign bicycles.
All variables are represented as deviations from their average during 1922–1935.

sumption one should remember that, because of the gaps in the available statistics, it was impossible to take into account such circumstances as irregularities in the demand for replacement caused by the stagnation of consumption during the war.

From formula (1) we can calculate that the elasticity of demand amounts to about 1.3.

TABLE 3

	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935
Purchasing power	99.6	101.0	98.4	97.4	102.8	102.1	104.6	109.5	108.8	102.4	96.6	91.9		
Price of foreign bicycles (in gld.)	39.4	34.2	35.7	44.8	35.5	30.3	39.8	53.9	58.5	48.1	30.0	29.0	24.8	23.7
Price of Dutch bicycles (in gld.)	88.9	75.4	69.4	62.6	61.1	54.9	53.7	52.9	51.6	47.0	40.3	32.1	29.7	28.2
Average price of all bicycles (in gld.)	59.1	53.8	59.6	61.5	59.1	52.8	53.2	52.9	51.6	47.0	40.0	32.0	29.6	28.1

TABLE 4

HOME PRODUCTION, IMPORTS, AND EXPORTS OF BICYCLES (IN THOUSANDS)

	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935
Home production	117	121	139	213	266	308	374	403	397	321	296	415	416	399
Imports	166	129	57.8	13.7	22.6	22.6	10.4	4.2	2.5	2.8	5.4	4.9	5.7	6.6
Exports	6.8	7.1	11.0	15.5	13.1	12.5	18.3	12.7	9.1	5.2	2.7	3.4	3.6	3.2

TABLE 5

IMPORTS OF BICYCLES FROM GREAT BRITAIN AND GERMANY

	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935
From Great Britain:														
Number of bicycles (thousands)	18.1	11.5	9.6	5.3	3.7	3.5	3.5	2.2	1.6	1.5	4.1	3.0	3.6	5.1
Average price (in gld.)	64.7	61.0	58.1	60.9	61.9	63.3	57.1	60.5	65.8	59.3	31.5	33.0	28.3	26.9
From Germany:														
Number of bicycles (thousands)	147	114	44	6.7	11.1	17.3	4.8	0.7	0.2	0.4	0.4	0.3	0.2	0.1
Average price (in gld.)	36.0	31.0	29.7	30.5	26.5	22.5	24.2	38.8	42.9	25.1	22.5	28.1	26.7	27.3

II. THE HOME PRODUCTION AND THE IMPORTS OF BICYCLES

The Dutch bicycle manufacturers have succeeded in making headway against foreign competition. Whereas in 1922 the greater part of the bicycles was imported, the share of imports in total consumption was less than 2 per cent after 1930 (Table 4). Further investigation showed that the interrelation between home production and imports can be chiefly explained by the prices of foreign and Dutch bicycles (Figure 3). This relation is shown by the following equation:

$$b = 87 + 0.77\,(P_B - 38) - 1.22\,(P_N - 53), \tag{2}$$

in which b is the home production as a percentage of total annual consumption, and P_N and P_B the average price of Dutch and foreign bicycles respectively.

The correlation coefficient was found to be 0.90.

This figure, however, may be affected by the circumstance that differences in quality could not be taken into account, as the available statistics do not give any information to this point. From the figures on the imports of bicycles from Great Britain and Germany, given in Table 5, it follows that the differences in quality are probably not insignificant.[3]

At the base of the development of the Dutch bicycle industry there is a high degree of rationalization, consisting principally in standardization. In connection with the fact that the necessary semi-finished articles and parts are now also made in Holland, this rationalization has made possible an important price reduction.

III. THE EXPORTS OF BICYCLES

The exports of bicycles from the Netherlands are not very important. They have declined since 1928 and now amount to less than 1 percent of the total home production. The main part is destined to the Netherlands East Indies and has to compete there with imports from Great Britain and for some years especially from Japan. On the other hand the imports of German bicycles have declined.

J. B. D. Derksen
A. Rombouts

The Hague

[3] It may be observed that a curvilinear regression would have given a closer relationship.

COMMENTS ON MILLIKAN'S REVIEW OF PARETO'S SOCIOLOGY

By Volrico Travaglini

Apropos the English translation of the *Trattato di Sociologia Generale,* of Vilfredo Pareto, Max Millikan takes occasion to present to the readers of Econometrica his interpretation of Pareto's sociological work.[1]

According to the explicit statement of the author, this interpretation should enable the ignorant or ingenuously enthusiastic reader of the *Trattato* to "approach the book with a full realization of its inadequacies," and should adequately prepare him to discover "the true magnitude of Pareto's contribution" amid "the welter of inconsistency and superfluity which envelopes the very meaty substance" of Pareto's thesis.

Millikan's essay in its totality seems to me open to discussion, because the *Trattato* is not considered as a sociological work, which it is in its essence, but as a philosophical one. The first part of the essay (Pareto as Sociologist), which should have been the fundamental one, has really only a purely expository character, with some important and a few absolutely incidental and negligible evaluations.

Millikan sees in the *Trattato* a late and mistaken exaltation of the inductive method rather than a serious and worthy attempt at the theoretical arrangement of the social forces and processes. It is on this conception, rather than on the exposition and critical evaluation of Pareto's sociological theories, that the author has based his arguments and conclusions, which are substantially unfavorable to Pareto's sociological work.

But also as regards the second part of Millikan's essay (Pareto as Scientist) I do not entirely agree with the author. The inquiry concerning the scientific method applied by Pareto and as to the results he obtains through the application of this method could have been, to my mind, efficacious, and consequently could have made the interpretation and evaluation of the sociological work of Pareto really useful, only if the author:

1. Had actually proved the errors of Pareto's sociological theories as a consequence of the procedure followed;
2. Had taken a clear stand between realism and idealism.

Concerning the first point, one may observe that Pareto, while he has frequently affirmed the superiority and necessity of the inductive

[1] Max Millikan, "Pareto's Sociology," Econometrica, Vol. 4, Oct., 1936, pp. 324–337.

method, has also clearly showed his sympathy with the deductive method. He has applied, indeed, both procedures.

To find a confirmation for the orthodox utility of the methodological affirmations of Pareto, which Millikan criticizes, one could express for the *n*th time the greatest admiration and warmest praise for the inductive stage of scientific development, which has taken the place of and surpassed, or better, which is thought to have taken the place of and surpassed, from the Renaissance on, the purely deductive stage, dear to the classic and mediaeval world. But it would be out of place to insist today on the old and exploited argument, that it is evidently impossible to think of the priority or the superiority either of induction or of deduction. Observation and experiment cannot go beyond a priori deduction. Induction is always based on some universal principle, interpreted sometimes as the uniformity of nature, at other times as universal causality, etc. This principle, which constitutes the basis of induction, can evidently not rest on induction. Therefore, strictly, one cannot speak of the priority or superiority of induction or deduction. Both methods are at the same time the basis for Science, which unfortunately cannot possibly reach a complete fusion of the two.[2]

Regarding the second point, I notice that Millikan bases his theory, in several remarkable sections of the second part of his essay, on the authority of Russell, Cohen, Whitehead, etc., that is to say on the authority of some of the most brilliant exponents of contemporary realism.

To what extent are we justified in believing that he follows their doctrines? We do not know exactly, but contemporary realism, in its different and sometimes conflicting currents, certainly has in its substance, together with the traditional English empiricism which has evolved from Locke to Stuart Mill, that methodological procedure condemned by Pareto, which reduces the totality and complexity of reality into its elemental stages and denies to thought all autonomy and all creative capacity. Contemporary realism, especially in some of its tendencies (more in Russell than in Whitehead and Alexander), reducing the importance of perceptive fact to the advantage of knowledge, tries to give a cold, objective, total vision of reality and consequently unconsciously destroys a certain pre-eminence of spiritual fact over natural fact, which the classic empiricist implicitly admitted as long as it stopped at sensation, that is at an absolutely subjective element. Realism is generally concerned with reducing the subject to a particular instance of natural objectivity, which in substance submerges and nullifies the subject itself. How does Millikan reconcile these and other

[2] Cf. R. G. Collingwood, *Speculum Mentis or the Map of Knowledge*, 1924, p. 177–180. V. Travaglini, *Gli schemi teorici del movimento della popolazione*, Padova, 1929, p. 298, etc.

affirmations of realism with the assertion contained in the second part of his essay?

Only the new realism of idealistic tendencies of the second Whitehead,[3] who revises and surpasses in his latest works all his first and prominent realistic conclusions, could be reconciled with Millikan's arguments and could at the same time open the door to that interpretation of the sociological work of Pareto, which seems the best to me. In his work, *Process and Reality*, Whitehead assigns a new and remarkable role to feeling in the structure of his system. In such a work the conception of an impartial relation between mind and things disappears. Such a conception is a chimera, for which reason every "sociology" is an imperfect construction: necessarily imperfect *in re* first and more than imperfect because of the eventual deficiencies of particular cultivators or by reason of the particular methods which they follow. Further, it is imperfect because the materials of perception and historical element on which it rests, rather than as changeable, complex and fluid, which they really are, or, rather, as we actually live and feel them to be, become, through his elaboration, ingenuously conceived as existing for themselves, that is to say, independently of the thought which places them, in their own absolute necessity and mechanism. Finally, it is imperfect because, for such a theory, those materials and those elements are necessarily simplified and more or less unconsciously immobilized and solidified. They are reduced to the conventional forms of types and classes; determined one for the other in their various parts; chained one to the other for the succession of concepts and the series of causes; reduced to the rigid frames of formulas and schemata.

Thus life and history, and particularly social process, lose their warm, inexhaustible, and unforeseen variety and complexities of forms and events. They are reduced to mere mechanism and determinism. They become a complex of necessary and continuous relations, an identical whole of determinations. Development, changed into categories of abstract thought, is reduced to the static unity of *relations* and *laws*, which destroy all spatial and temporal contingencies, and for which reality, our reality, is *ab aeterno* immobilized, and should delusively be known only once and for all.

VOLRICO TRAVAGLINI

Royal University of Catania
Catania, Italy

[3] A. N. Whitehead, *Process and Reality. An Essay in Cosmology*, Cambridge, 1929.

FINAL ANNOUNCEMENT OF THE EUROPEAN MEETING, SEPTEMBER 11–15, 1937

The European Meeting of the Econometric Society for 1937 will be held at Annecy, Haute Savoie, France (not far from Geneva), in the Hotel des Trésoms et de la Forêt from September 11 to 15, 1937. The meeting will begin at dinner on September 11; papers will be read on the following days.

Rooms with full board are available at Fr. 45 to Fr. 55, depending on location; a limited number of smaller rooms with simpler meals is also available at Fr. 30 to Fr. 35. Applications for rooms should be made as soon as possible to J. Tinbergen, Société des Nations, Geneva; no accommodations can be guaranteed to those applying after August 25. Further details of the program and a prospectus showing the location of the hotel will be sent to those who inform J. Tinbergen of their intention to attend the meeting. Those who notify the hotel directly of the exact date and hour of their arrival will be met at the Annecy station by a motorbus.

EDITOR'S NOTE

An extensive report of the papers and discussion at the Sixth European Meeting of the Econometric Society, Oxford, 1936, has been prepared, but because of lack of space must be deferred until the October issue.

THE NATURE OF REGRESSION FUNCTIONS IN THE CORRELATION ANALYSIS OF TIME SERIES

By Herbert E. Jones

I. INTRODUCTION

Though the general theory of correlation has been extensively developed since its introduction by Galton and Pearson, most of this development has been in that field of statistics which cannot be adapted, without extensive modification, to the analysis of time series.[1] One of the difficulties in correlating time series is the question of the linearity of the regression function. Pearson's "eta" test for nonlinearity and other tests that depend on the frequency distributions of the data are, as we shall see, not applicable to time series. For this reason most of the literature on time correlation assumes linear regression, but there are a great many cases when the regression is far from linear and our ordinary methods prove inadequate.

It is the purpose of this paper to investigate the nature of the regression function by a new method to determine, not only the existence of nonlinearity, but also the general form or shape of the regression function. It will also show the advantages of graphical analysis in the study of lag and lag correlation.[2]

II. THEORETICAL FORMULATION

1. Classification of Time Series. An examination of the form or shape of time series will show that the cyclic movements are not, in general, symmetrical. Some series will have sharp crests or flat troughs, others are skewed to the right or left; a variety of combinations is possible. These shapes, however, are often quite consistent within any specific series over long periods of time but may vary widely between different

[1] For a good discussion of these problems see C. F. Roos, "Annual Survey of Statistical Technique: The Correlation and Analysis of Time Series—Part II", Econometrica, Vol. 4, Oct., 1936, pp. 368–389.

[2] R. Frisch has made an extensive investigation of the relations between the scatter diagram and linear correlation in his very interesting papers, "Correlation and Scatter in Statistical Variables," *Nordic Statistical Journal*, Vol. 1, 1928, pp. 33–102 and *Statistical Confluence Analysis by Means of Complete Regression Systems*, Oslo, 1934. Frisch's derivations are quite general and are not limited to data ordered in time, or by the number of variables included. He establishes various properties of his coefficient of scatter which is a measure of the closeness of a set of observed quantities to linear dependency, and he is chiefly interested in linear correlation from the point of view of linear transformations. The present paper approaches the problem from a different point of view and attempts to determine the nonlinearity of the scatter for time series in contrast to Frisch's measure of departure from linearity for any kind of a statistical population.

series; and, as we show later, only when series of the same shape are correlated is the relation between them linear. Therefore, in general, the correlation will be nonlinear, but when the difference in shape is small, a linear regression can be used.[3]

Time series in our analysis will be classified under two general criteria of form, one, a criterion of "steepness," the other, a criterion of "skewness." To segregate the cycles under consideration it is necessary to eliminate the trend and for the present it is assumed that this has been done. Since the upper part of the cycles may have quite a different shape from the lower, the cycles are divided into two parts. That half of the cycle lying above the trend we call the positive half cycle, that lying below the trend, the negative half cycle.

2. Effect of Steepness. To measure the "steepness" a criterion is needed that is invariant to the scale factor. This independence of scale factor is quite necessary since it is not the common conception of "steepness" (the ratio between the vertical and horizontal distance) that is important, but the "flatness" or "sharpness" of the cycle, especially in the extremes. The coefficient of variation satisfies these conditions very well, being invariant to the scale and quite sensitive to the shape of the data.[4] This criterion we designate by v, using v' to denote the coefficient for the positive half cycle, v'' the coefficient for the negative half. Although v is often used as a percentage it will be used throughout this paper as the ordinary ratio,

$$v=\sigma_x/\bar{x}, \tag{1}$$

where σ_x is the standard deviation of the observed points lying on *one side* of the trend and $\bar{x}$ the mean of these points.[5] This criterion will be relatively small for flat curves, growing larger as the steepness increases, with limits of the order of 0.2 for very flat curves up to 0.9 for very steep curves.

Figure 1 shows the effect of steepness on the form of the regression curve. Figure 1-A shows two hypothetical series: the X curve is rela-

[3] Whether or not this difference is significant will depend on the sampling errors of the criteria. Unfortunately, we have not been able to determine these sampling errors as yet and for that reason the method is still in a preliminary stage.

[4] L. Hersch in his paper, "Essai sur les variations periodiques et leur mensuration," *Metron*, Vol. 12, No. 1, 1934, pp. 3–184, defines various indexes or criteria for classifying periodic series. He includes the coefficient of variation but does not mention that it will measure the steepness of the cycle. His index for this characteristic is much more complicated but no more efficient.

[5] The question naturally arises whether all the points above the trend can be grouped together in equation (1) or if each cycle should be investigated separately. This question will be discussed later when we consider the practical applications; for the present one cycle only will be studied.

tively flat while the Y curve is relatively steep. Figure 1-B shows the "regression curve" of the two series which is actually decidedly non-linear though one would say offhand that the correlation was high.

A more detailed study of the X and Y series shows that as the curves leave the point t_0 the slope of the X curve is much greater than that of the Y curve but this relation is reversed before they reach their maximum point E at t_3. The associated pairs of points, (x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3), are plotted in Figure 1-B, and shown as the line 0–3

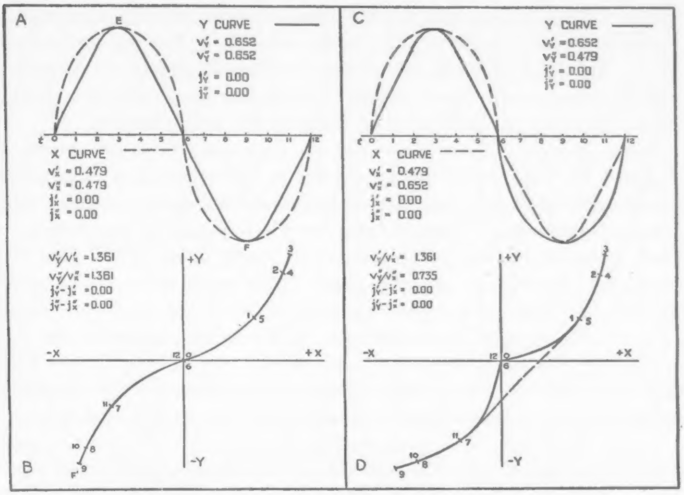


Figure 1.—Hypothetical Series showing the effect of the coefficient of curvature on the regression function.

with a decided curvature towards the Y-axis. As the points travel from t_3 to t_6 in Figure 1-A the curve 0–3 is retraced from 3 to 6 since the X and Y curves are symmetrical about a vertical axis. In a similar manner as the points move from t_6 to t_9 we trace the curve 6–9. If the designations of the two series were reversed, a curvature towards the X-axis in Figure 1-B would result. Also, if the two cycles had been negatively correlated, that is, if the negative half of one cycle corresponded to the positive half of the other, the regression curve would lie in the 2nd and 4th quadrants, and the curvature would again be toward the axis of the steeper curve. Had the two cycles had the same steepness, their regression curve would have been a straight line through the origin.

This curvature in the regression function is due to the *difference* in

the "steepness" of the two cycles, and the regression tends to curve towards the axis of the steeper curve. It is not the absolute values of criteria of steepness that determines this curvature but some comparative relation between the two.

We may define a coefficient of curvature as v_y/v_x; as $v_y/v_x \gtreqless 1$, the slope of the the regression function respectively increases, is constant, or decreases from the origin.

To get a true picture of the regression function the positive half cycles and negative half cycles must be investigated separately. In Figure 1-C are shown two hypothetical series whose criteria of steepness are the same as before, but in this case the Y curve is the steeper in the positive half while the X curve is the steeper in the negative half. In other words, the coefficient of curvature is greater than unity in the 1st quadrant and less than unity in the 3rd quadrant.

The coefficients of curvature and the regression curve are shown in Figure 1-D. This shows a salient point at the origin, but in practice the means of the two series are not, in general, coincident in time,[6] and the actual regression curve follows some other path in the neighborhood of the origin such as the dotted line shown in the figure. So if the coefficient of curvature is either greater than unity or less than unity for both half cycles the regression is S-shaped;[7] if the coefficient shifts from a value greater than unity to a value less than unity for the different half cycles the regression is more like a hyperbola. In the first case the relationship could be well approximated by a cubic parabola while the second case a quadratic equation will usually suffice.[8]

3. Lag Hysteresis. In the above illustrations both the positive and negative half cycles were symmetrical about vertical axes through their centroid, that is, they were not skewed. In actual time series this symmetry is rarely present so the question of skewness and its effect

[6] Karl Pearson remarks on this feature in his classic paper, "Skew Correlation and Non-Linear Regression," *Drapers' Company Research Memoirs*, Biometric Series II, 1905, p. 29. Pearson writes: " . . . the regression curve does not pass through the mean of the two characters, . . . an individual with the mean of one character will most probably not have the mean of a second character. This is an important result, which follows at once for nearly all types of skew correlations."

[7] K. Pearson also remarks on this point, *loc. cit.*, p. 28. He says: "In general the regression line will be part of an S-shaped curve." He gives a reference to a typical S-shaped regression curve but makes no explanation of this characteristic. The data Pearson usually worked with were not time series and the above analysis would not apply in his case.

[8] The above analysis is somewhat related to Ragnar Frisch's theory of changing harmonics, since Frisch's operator could be considered as a criterion of shape. See his "Changing Harmonics and Other General Types of Components in Empirical Series," *Skandinavisk Aktuarietidskrift*, 1928, pp. 220–236.

on the regression function must be studied. Also in these examples the hypothetical series were perfectly in phase, that is, there was no "lag" between the series. Although at first thought there may not seem to be any connection between the concepts of skewness and lag, it will be seen presently that they are very closely related, being, in fact, the two components of what we shall call *economic hysteresis.*[9]

In the ordinary language of time-series analysis this factor has usually been termed "lag," but here it is necessary to use hysteresis with its broader concept meaning the retardation of the effect of a changing

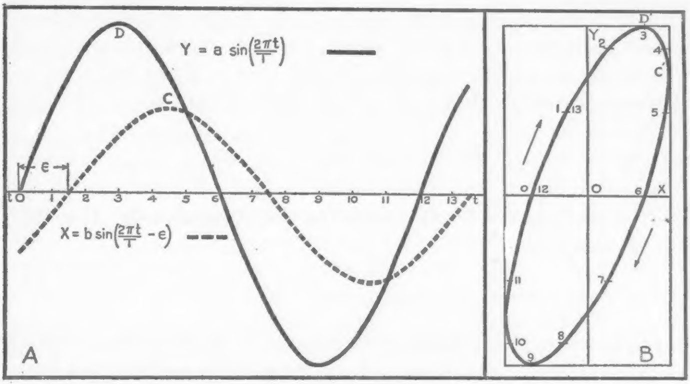


FIGURE 2.—Effect of lag hysteresis in the correlation of two sine curves

force. The ordinary lag, or what will hereafter be called "lag hysteresis," is but one part of the general hysteresis, the other part being the "skew hysteresis." The lag hysteresis can be corrected by the ordinary means of changing the phase, or "lagging" one of the series. The skew hysteresis cannot be corrected in this manner, as is shown later.

For simplicity two sine curves will be used for hypothetical series to illustrate the effect of lag hysteresis on the regression function. Let the X curve be defined by

$$x = b \sin\left(\frac{2\pi t}{T} - \epsilon\right) \tag{2}$$

[9] Hysteresis was first introduced into econometrics by C. F. Roos, "A Mathematical Theory of Competition," *American Journal of Mathematics*, Vol. 47, 1925, p. 173, who showed that the integral derived in the demand equation when past prices are included leads to the basic equation of hysteresis as used in physics. This is a very appropriate term for econometrics. It comes from the Greek, meaning to lag or be behind, and is used in physics as the lagging or retardation of the effect of a changing force. This is exactly what is meant by "lag" in the economic analysis of time series.

and the Y curve by

$$y = a \sin\left(\frac{2\pi t}{T}\right), \tag{3}$$

where ϵ is the phase angle or "lag."

These curves are illustrated graphically in Figure 2-A. In this simple case the functional relationship between x and y can easily be shown to be

$$a^2x^2 - 2abxy \cos \epsilon + b^2y^2 = a^2b^2 \sin^2 \epsilon. \tag{4}$$

Eq. (4) is an ellipse with its center at the origin of the co-ordinate system and with its axis inclined with respect to the X-axis at an angle whose tangent is a function of ϵ. If $\epsilon = 0$, that is, if the two series are in phase, the ellipse defined by (4) degenerates into the straight line

$$y = \frac{a}{b}x, \tag{5}$$

and the coefficient of linear correlation would equal unity. If $\epsilon = 90°$, Eq. (4) becomes

$$\frac{x^2}{b^2} + \frac{y^2}{a^2} = 1, \tag{6}$$

an ellipse whose axes coincide with the axes of the co-ordinate system and whose semi-axes equal a and b. In this case the coefficient of linear correlation would equal zero. When ϵ has any value between these two limits,[10] Eq. (4) defines some ellipse such as that shown in Figure 2-B. As the associated points, (x_1, y_1), (x_2, y_2), $\cdots$ (x_{12}, y_{12}), proceed with time, the point on the Y series reaches its maximum, D, at t_3 while the X point reaches its maximum, C, at $t_{4.5}$. Therefore, as these points travel with time they trace the ellipse shown in Figure 2-B. Point D, the maximum co-ordinate in the Y direction, will precede point C, the maximum co-ordinate in the X direction; in other words, the points will trace the ellipse in a clockwise direction. By the same analysis, if the X series had preceded the Y series, the points would have traveled in a counterclockwise direction. If the two series were negatively correlated, the ellipse would lie in the 2nd and 4th quadrants and, in that case, a clockwise rotation of the points would indicate that the X series preceded the Y series.

It is also worth noting that the points C and D are the points on the ellipse tangent to the circumscribed rectangle whose sides are $2a$ and $2b$. Therefore, when the associated points of series X and Y are plotted in a scatter diagram they form an open loop of elliptical shape

[10] If $\epsilon > 90°$ it could be considered as negative and less than $90°$.

and the difference in time between the point of horizontal tangency and the point of vertical tangency indicates the "lag."

Economic series are not sinusoidal nor are they smooth curves. But a scatter diagram of two cyclical series will, if the series are not in phase, show a loop construction. If successive points in the scatter diagram are connected, a rough estimate of the lag can be made by computing the time necessary to go from a maximum in the Y direction to the maximum in the X direction, that is, from the point of horizontal tangency to the point of vertical tangency. When this lag is corrected by changing the phase of one of the series, the points in the scatter diagram will tend to concentrate about a curve which is the regression curve. In this manner the "lag hysteresis" may be corrected.

The above analysis can be stated in the general theorem: When successive points in a scatter diagram are connected, loops indicate the existence of "lag hysteresis." If the series are positively correlated and successive points rotate in a clockwise direction, the series used for the Y variable of the scatter diagram precedes; a counterclockwise rotation indicates that the series used for the X variable precedes. If the series are negatively correlated, these rotations are reversed. This "lag hysteresis" can be corrected by lagging the items of the two series in the necessary manner until the rotational direction becomes random or is reversed. If the lagging of one more item causes a consistent reversal in the direction of rotation, the optimum lag lies between the preceding lag and the lag causing the reversal.[11]

4. Skew Hysteresis. The above analysis has shown the effect of "lag hysteresis" on the regression function and shows that this effect can be corrected by the ordinary method of lag correlation. But the hysteresis caused by skewness cannot be so easily corrected.

The criterion of skewness which we have chosen, after some investigation as to its necessary attributes, is the ratio of the time interval, between the vertical axis through the centroid of the area under the half cycle and the midpoint of the half cycle, to the length of the half cycle.[12] It can easily be computed by the following formula:

$$j_x = \left[\frac{\sum_{i=a}^{i=b} x_i t_i}{\sum_{i=a}^{i=b} x_i} - \left(\frac{t_a + t_b}{2} \right) \right] \Big/ (t_b - t_a), \tag{7}$$

[11] R. von Huhn has developed another method by which the lag can be measured graphically, but it is more complicated tha the above analysis; see his article, "Secondary Curves As a Measure of the Lag or Phase Difference Between Two Primary Curves," *Jour. Am. Stat. Ass.*, Vol. 28, 1933, pp. 312–327.

[12] L. Hersch, *op. cit.*, has what he calls an index of symmetry but it does not

where t_a and t_b, respectively, refer to the times at the beginning and the end of the half cycle. From Eq. (7) it is seen that if the centroid lies to the right of the midpoint the skewness is positive; if it lies to the left, it is negative. If we neglect the denominator, the skewness will be in terms of the time interval, and in many cases this is preferable.

The manner in which the regression function is affected by skewness is shown in Figure 3. The positive half cycle of the Y curve in Figure

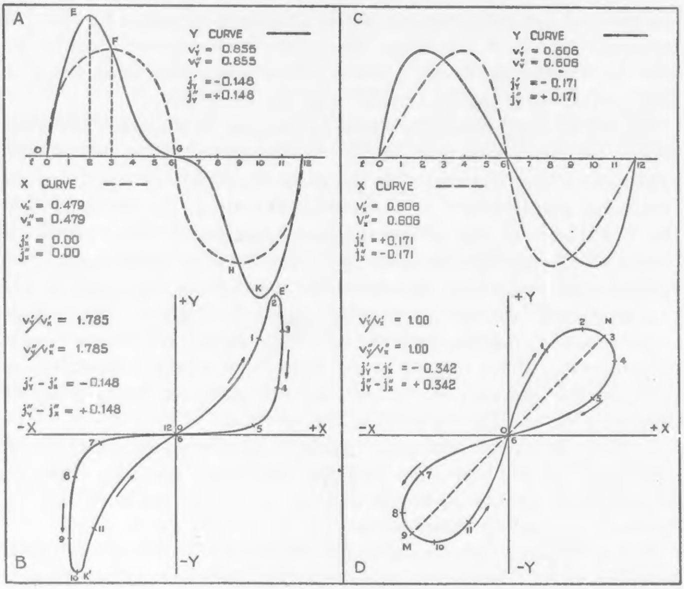


FIGURE 3.—Hypothetical Series showing the effect of skew hysteresis on the regression function.

3-A is negatively skewed, while the negative half cycle is positively skewed. In actual time series it is usually found that the criteria of skewness of the positive and negative half cycles are of opposite sign. With the same notation as in the case of "steepness," j'_x will indicate the skewness of the positive half cycle of the X curve and j''_x will indicate the skewness of the negative half cycle. In Figure 3-A the X curve

fit the needs of the present problem and entails more computation. Since nothing is known about the standard errors of Hersch's indexes there seems to be no advantage in using his criteria in the present analysis.

has no skewness and the skewness of the Y curve is $j'_y = -0.148$ and $j''_y = 0.148$; the criteria of steepness are,

$$v_y = 0.855 \quad \text{and} \quad v_x = 0.479.$$

In Figure 3-B is shown the functional relationship between the X curve and the Y curve. As the associated points, (x_0, y_0), (x_1, y_1), $\cdots$ (x_{12}, y_{12}), travel from t_0 to t_{12} in Figure 3-A, their regression function traces the curve 0, 1, 2, 3, $\cdots$, 12 as shown by the numbers in Figure 3-B. The Y curve reaches its maximum point E at t_2 while the X curve reaches its maximum F at a later time t_3. In other words, because of the negative skewness in the Y curve, the Y curve reaches its maximum point first. This causes a loop to be formed in the regression function as is shown by the points 0, 1, 2, $\cdots$ 6; the Y co-ordinates decrease after reaching 2 while the X co-ordinates continue to increase until they reach 3. The skewness of the Y curve is positive in the negative half cycle, and the regression curve of the 3rd quadrant has a loop in the opposite direction because the X curve is first to reach its maximum.

The coefficient of curvature is

$$v'_y/v'_x = v''_y/v''_x = \frac{0.855}{0.479} = 1.785,$$

which, according to the previous concepts, indicates a curvature of the regression function towards the Y-axis in both the 1st and 3rd quadrants. This is quite apparent in Figure 3-B. Figure 3-D shows the regression curve of two hypothetical series whose skewness is equal but opposite in sign and for which the coefficient of curvature indicates linearity. Their regression function, however, is far from linear because of the skewness of the curves. The skewness causes the figure-8-shaped loop, and the linearity of the coefficient of curvature is shown in the symmetry of the loop about the straight line MN.

Again, as in the consideration of "steepness," it is not the absolute value of the criterion of skewness which is important but its value in relation to the other series with which it is being correlated. If the skewnesses in two series were equal in value and sign, there would be no loop in the regression curve, and so the difference of the two skewness criteria is a good measure of this characteristic. Therefore, we define the coefficient of "skew hysteresis" by

$$J = j_y - j_x. \tag{8}$$

It must be computed for both the positive and negative half cycles.

A study of Figure 3 shows that when the maximum point of the Y curve precedes the maximum of the X curve the successive points in the regression function travel in a clockwise direction; while if the

maximum of the X curve precedes, the points travel in a counterclockwise direction. This condition will be indicated in the coefficient of "skew hysteresis" by its sign, a negative value indicating that the Y curve precedes, and a positive value indicating that the X curve precedes. Both illustrations in Figure 3 are, however, examples of positive correlation. A similar study of negative correlation shows that the loops will lie in the 2nd and 4th quadrants and the coefficient for the 2nd quadrant will be

$$J = j'_y - j''_x, \tag{9}$$

and for the 4th quadrant

$$J = j''_y - j'_x.$$

The directions of the loops are reversed from what they were in the case of positive correlation. That is, in the case of negative correlation, if the Y curve reaches its maximum before the X curve, the loop is formed in a counterclockwise direction, and if the X curve precedes, the loop is clockwise.

This can be stated in the following theorem: If the coefficient of "skew hysteresis" is different from zero a loop will be formed in the regression curve. A negative coefficient indicates that the Y curve precedes; a positive coefficient indicates that the X curve precedes. If the series are correlated positively, a negative coefficient indicates a clockwise rotation and a positive coefficient indicates a counterclockwise rotation; and vice versa if the series are correlated negatively.

From Figure 3-B it is seen that the portion of the regression curve, 10, 11, 0, 1, 2, can be approximated very closely by a straight line, but certainly no straight line can approximate the portion 2, 3, on to 9, and 10. On the other hand an S-shaped curve approximates both branches of this curve quite well. In the example shown in Figure 3-D, one would have to be satisfied by the line MN if a linear regression function were desired. A much closer approximation, however, can be made by using two regression functions, one for the upward swing of the cycle and one for the downward swing. That is, if the coefficient of skew hysteresis is large, we must use two regression functions.

The question arises as to whether or not this "skew hysteresis" can be corrected in a manner similar to the correction for "lag hysteresis." A brief study of the series given in Figure 3 shows this to be impossible.

In the case of "lag hysteresis" the regression was simplified by lagging one of the series until the two were "in phase." While this method is applicable in the case where the two series have equal skewness, it is easily seen that new difficulties arise when this condition is not satisfied. In the series shown in Figures 3-A and 3-C there are *three*

ways in which the two series may be said to be "in phase." The points of inflection may be synchronized as shown in Figure 3, or the maximum points may be synchronized, or the minimum points. But in the last two cases the loop is only emphasized as shown in Figure 4. Figure 4-A shows the same series as given in Figure 3-A, with the exceptions that the designations of the X and Y curve are interchanged and that the maximum points have been synchronized. The resulting regression

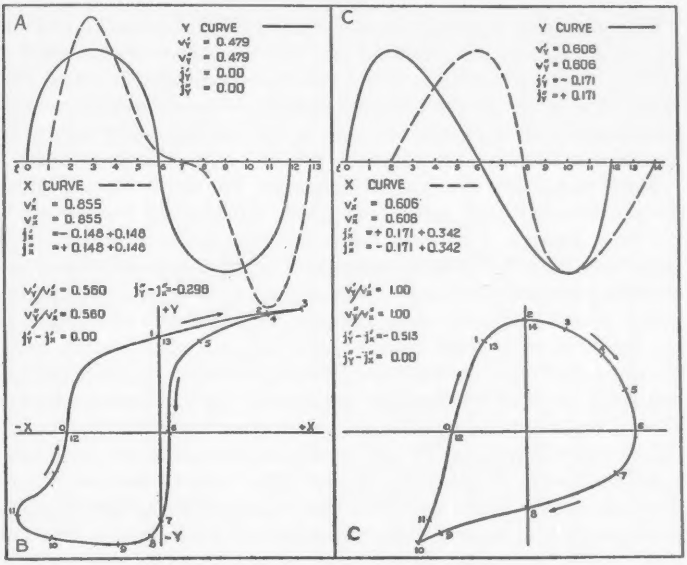


FIGURE 4.—Distortion in the regression function when the attempt is made to correct skew hysteresis.

function is shown in Figure 4-B. In other words, a lag of 0.148 has been introduced which cancels the "skew hysteresis" in the positive half cycle but doubles it in the negative half. The resulting regression curve is quite complicated and it would be impossible to represent it by a simple function. Also such a relationship as shown would have doubtful value for giving the relation between the X and Y series. Again, in Figures 4-C and 4-D are shown the effect of synchronizing the minimum points of the series of Figure 3-C. The loop has been increased, and it must be concluded that these induced lags instead of reducing have exaggerated the effect of skewness.

Therefore, "skew hysteresis" cannot be entirely corrected by simple means. The optimum condition will be found when a lag is introduced such that the coefficients of skew hysteresis are equal but of opposite sign for the positive and negative half cycles. If the resulting coefficient is large, two regression functions will be necessary to show the interrelationship, for the upward and downward swings of the cycle respectively.

III. PRACTICAL APPLICATIONS

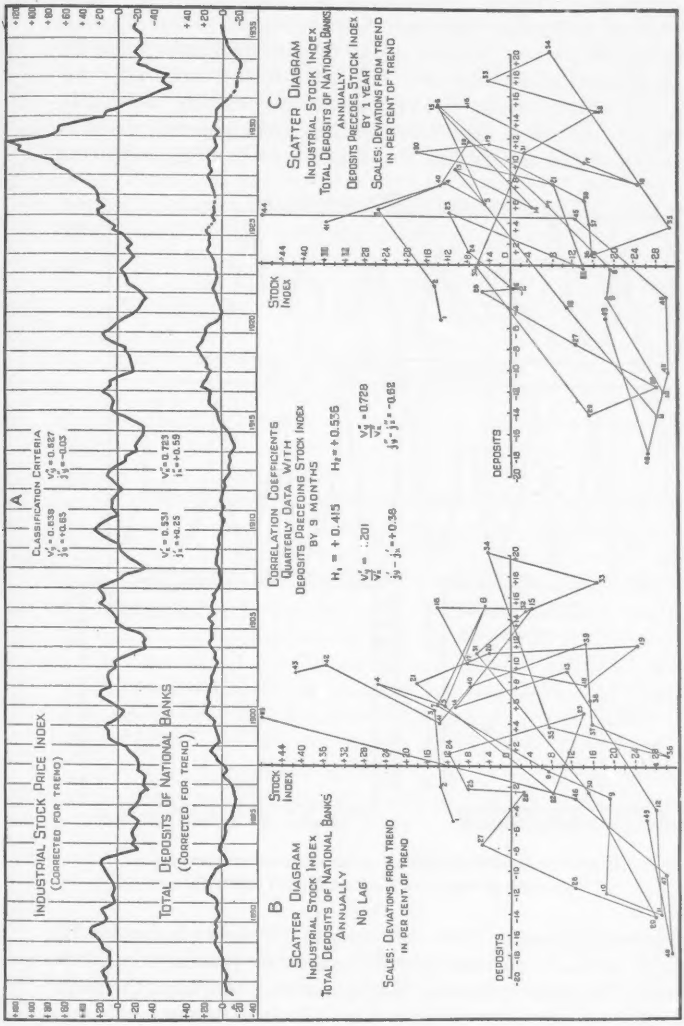
The preceding section was based on studies of smooth, periodic curves. Such an analysis would have little value in econometrics if it could not be applied rather easily and with advantage to actual economic time series. In the following section three applications are discussed and it is seen that not only is the analysis easily made but also certain definite advantages are gained over ordinary methods.

Figure 5-A shows two typical time series. The upper series is an index of industrial stock prices from 1896 to 1936;[13] the lower series is the total deposits of national banks covering the same period. Both series have been corrected for trends and are shown as deviations from trend in percentage of trend.[14] In these graphs the wide divergence between actual economic series and the hypothetical curves used in Section II is easily seen. But the same analysis is applicable. After removing the trend, the coefficient of steepness was computed for each half cycle. In every case but one in the stock price index the criterion showed that the positive half cycle was steeper than either the preceding or succeeding negative half cycle; the deposit series gave consistently steeper negative half cycles. The value of the criterion of steepness for the whole series was then taken as the arithmetic mean of the individual values. These values are shown in Figure 5-A. The criterion of skewness was then determined for each series, the computation being in terms of the time interval and not in terms of the half cycle. This method of computing the skewness is better in actual practice since the values will indicate directly in time units the lag needed to make the coefficients of skew hysteresis equal and opposite in sign.

Next the scatter diagrams were drawn. The original data were in quarterly figures which means 196 pairs of values. A scatter diagram of so many points would hopelessly confuse the pattern, and hence

[13] Cowles Commission—Standard Statistics, Index of Industrial Stock Prices.

[14] These series are given here only as an example of the above method of analysis. An explanation of a probable economic relation between these series was given by Alfred Cowles 3rd in his paper, "Effects of Building Activity and Other Factors on Security Prices," presented before the Econometric Society at St. Louis, January 3, 1936; see abstract in ECONOMETRICA, Vol. 4, Apr., 1936, pp. 190–191.



Deviations From Trend (per cent of trend)

FIGURE 5.—Scatter diagrams in the correlation analysis of Industrial Stock Price Index and Total Deposits of National Banks.

annual averages were used. These are shown plotted, without lag, in Figure 5-B. At first glance this figure seems to be just a maze of points with no distinct pattern, because of the random fluctuations and the presence of other movements still appearing in the series after the removal of the trend which tend to obliterate any definite pattern. If, however, the pattern is traced successively from point to point in Figure 5-B it will be noticed that the points travel in a very consistent

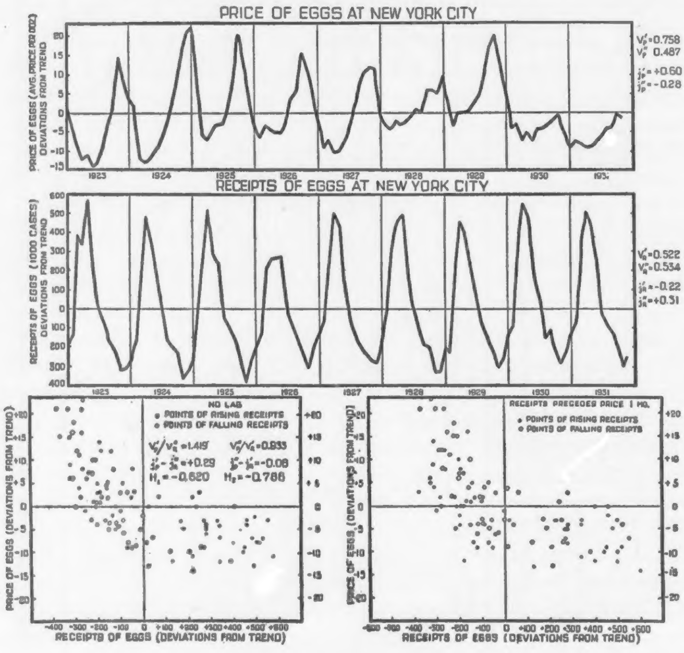


FIGURE 6.—Scatter diagrams in the correlation analysis of Price and Receipts of Eggs at New York City.

counterclockwise direction. This indicates, as was shown in Section II, that the X series, in this case the deposit series, precedes the stock price index. The loops indicate a "lag hysteresis," but, since the loops are not wide and are quite variable, it is difficult to determine the exact lag. Therefore, the items are lagged 1 unit, with deposits preceding, and the scatter diagram of Figure 5-C drawn. In this case when the pattern is traced from point to point, persistent loops are formed in a *clockwise* direction, indicating that the lag introduced was too large.

This shows that the optimum lag lies somewhere between 0 and 1 year with deposits preceding stock prices. Ordinary lag correlations give the maximum value when deposits lead stock prices by 9 months. In a lag-correlation analysis there is, however, a question as to the significance of the difference between the correlation for 0 lag and 9-months lag. The ordinary tests of significance fail when we make any reasonable assumption regarding the number of independent observations. But the scatter diagram for 0 lag shows such a consistent pattern of loops with a *counterclockwise* rotation that the existence of a significant lag is highly substantiated.

The coefficients of curvature and skew hysteresis are computed next. These are shown in Figure 5. The coefficients of curvature show indication of curvature towards the Y-axis in the 1st quadrant and towards the X-axis in the 3rd quadrant. The coefficients of skew hysteresis are very small when it is remembered that they are in units of 3 months. That is, less than 3 months would be necessary to equalize them but it is not possible to lag the items less than one unit and, since they are small, one regression function should suffice. Since the regression function curves toward the Y- and the X-axis it will be of the general type shown in Figure 1-D and can be represented by a second-degree polynomial. To determine whether a second-degree polynomial will fit the points better than a linear function, a generalized coefficient of correlation will be used. This was derived by Neyman[15] and is denoted by H_n where n is the degree of the polynomial. H_1 will be the same as the ordinary coefficient of correlation. In this problem $H_1=0.415$ while $H_2=0.536$. Although the usual test of significance of these values is lacking, because of our lack of knowledge about the independence of observations, the use of the second-degree polynomial seems to show a significant increase in correlation.

Another interesting example is given in Figure 6. Here are shown the price of eggs and the receipts of eggs in New York City from 1923 to 1932. One can see at once that the series are highly correlated nega-

[15] J. Neyman, "Further Notes on Non-Linear Regression," *Biometrika*, Vol. 18, 1926, pp. 257–262. In terms of discrete summation

$$H_n^2=\frac{1}{\nu_2\Delta_n}\begin{vmatrix} 0 & \lambda_0 & \lambda_1 & \cdots & \lambda_n \\ -\lambda_0 & \mu_0 & \mu_1 & \cdots & \mu_n \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ -\lambda_n & \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}, \text{ where } \Delta_n=\begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}$$

and $\mu_n=\sum x^n$, $\nu_n=\sum y^n$ and $\lambda_n=\sum x^n y$.
The variables x and y are taken as deviations from their means, so that

$$\mu_0=n,\ \lambda_0=\nu_1=\mu_1=0.$$

tively. The classification criteria, shown to the right of the series in the figure, indicate that the egg-price cycles have relatively steep positive half cycles with moderately flat negative half cycles. The skewness is positive for the positive parts and negative for the negative parts. Again the skewness is given in units of the time interval, one month in this case. In the supply series the criteria show relatively flat cycles in both parts of the cycle while the skewness is small and of opposite sign to that of the price series. This condition brings up a very interesting point. Since the series are negatively correlated, the positive half cycles of one series are compared with the negative half cycles of the other; and as the coefficient of skew hysteresis is the difference of the criteria of skewness, the skewnesses of the two series tend to compensate each other because they are of opposite sign. That is, when two series have skewnesses of opposite sign but are negatively correlated the skewnesses of the two series tend to balance each other.

The coefficients of curvature indicate a curvature toward the price axis in the 2nd quadrant and practically a linear relationship in the 4th quadrant. The skew hysteresis, as was said above, is very small and not significant. These values are shown in the figure. An analysis of the lag hysteresis is shown in the two scatter diagrams. In the lower left-hand diagram the points have been plotted without lag. The open circles are points of falling receipts while the solid circles are points of rising receipts. There is a slight tendency for the open circles to lie to the left of the solid, thus indicating a clockwise rotation. This means that receipts tend to precede prices. But the hysteresis loop is small and if we let receipts lead prices by one month, the second scatter diagram is obtained. This shows a slight indication of counterclockwise rotation which means that the lag introduced was too large, but, since the data are monthly figures, one has to be satisfied with either zero lag or with receipts leading by one month.

The general form of the regression function is indicated by the coefficients of curvature and is easily seen in the scatter diagram. A linear function gives the correlation coefficient H_1, or $r = -.620$, while if a second-degree polynomial is used $H_2 = -0.786$.

Figure 7 shows the relation between new-mortgage financing and the inhibiting influence of foreclosures.[16] In the two preceding illustrations the analysis covered several complete cycles. In this case the statistics are not available for even one complete cycle. Naturally the question arises whether the correlation from such meager data is significant.

[16] This illustration was taken from C. F. Roos, *NRA Economic Planning*, a monograph of the Cowles Commission for Research in Economics, Principia Press, Bloomington, Indiana, 1937, p. 425.

But in the analysis of economic data very often one does not have sufficient data for significant conclusions. He must do the best he can with the data available and be careful not to draw any startling conclusions from the results.

Since these series are so short we can only investigate one half cycle of each series, the negative half of the foreclosure series and the positive half of the new-mortgage financing series. The coefficients of

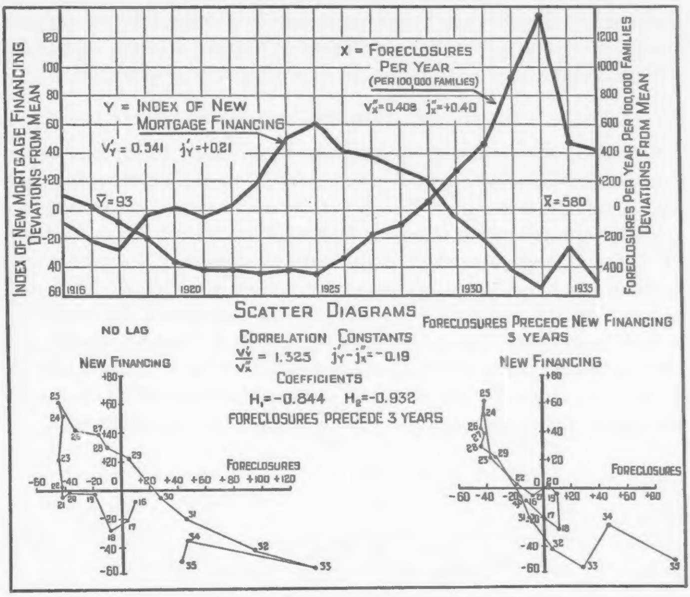


FIGURE 7.—Scatter diagrams in the correlation analysis of Foreclosures and New Financing.

curvature and the skew hysteresis are shown in Figure 7. The scatter diagram on the left shows the pattern formed by connecting successive points and is a good illustration of the "tear drop" pattern which was illustrated in Figure 4-D. This pattern is the result of the positive half of the foreclosure cycle and the negative half of the new financing cycle being in phase, while the other parts of the cycles are out of phase. A study of the loop formed in the 2nd quadrant shows a lag of from two to three years and as the pattern is formed with a clockwise rotation this indicates that foreclosures precede new mortgage

financing about two and one-half years. This lag is also confirmed by Roos' building study in which he found a lag of from $2\frac{1}{4}$ to $2\frac{1}{2}$ years by another form of analysis.[17] By introducing a lag of 3 years the second scatter diagram is obtained and it is seen that the hysteresis has been corrected about as well as can be expected. Although the first seven points (16 to 22 inclusive) show a reversal of the loop, the next 10 points (23 to 32 inclusive) show the typical crisscross pattern which indicates optimum correction. The coefficient of curvature shows a bending toward the Y-axis; also if a second-degree polynomial is used, the correlation coefficient is increased from -0.844 for a linear function to -0.932 for the second-degree polynomial.

IV. DISCUSSION OF THEORY

In Section II a criterion of steepness was defined and it was shown how the non-linearity of the regression function depended on the comparison of these criteria. The question might arise concerning the necessity of introducing this new coefficient since Karl Pearson's[18] correlation ratio measures the non-linearity of regression. Pearson's test, however, does not seem adequate for testing time series. It is well known that the correlation ratio is not independent of the classification of the data. In the analysis of those statistics which are distributed with respect to some physical characteristic and are grouped according to some logical class interval, the correlation ratio applies very well but is dependent partly on the classification. In such a case

$$\eta_{yx} = \frac{\sigma_{my}}{\sigma_y}, \tag{10}$$

where

$$\sigma_{my}^2 = \frac{\sum\left[n_x(\bar{y}_x - \bar{y})^2\right]}{N}, \tag{11}$$

and $\bar{y}_x$ is the mean of the y's in the xth array, and n_x is the number of y's in the array. In other words, the correlation ratio is the ratio of the standard deviation of the means of the arrays to the standard deviation of the y's in general. As the n_x's, the frequencies in the arrays, approach unity, σ_{my} approaches σ_y and η approaches unity. So, if there were just one observation of y for each observation of x and the x's could not be arrayed, η would equal unity regardless of the value of r and the difference between η and r would have no significance in regard to the linearity of the regression.

In the correlation of time series it is rarely possible to group the

[17] C. F. Roos, *Dynamic Economics*, Bloomington, 1935, p. 103, *et. seq.*
[18] *Op. cit.*, pp. 10, *et seq.*

data in class intervals. To use smooth periodic curves again as an illustration, consider the two time series in Figure 8. The X series has no secular trend while the trend in the Y series is very pronounced. If it is attempted to group the data in X arrays, such as the class interval centered at x_1, there will be a great variation in the values of y within the array and this variation is due mainly to the trend in Y. Also if the data are grouped in Y arrays such as Y_1, Y_2, Y_3, etc., the class interval must be very large to have any frequency within the array and a large portion of the cycle will be included in the interval, thereby masking the very effect that is being studied. Therefore, unless all movements except the one under investigation are eliminated, and

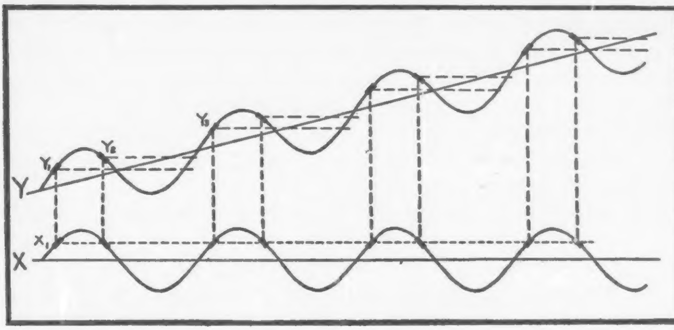


FIGURE 8.—Effect of trend in the classification of associated pairs of observations in time series.

unless the data can be classified without influencing the cyclic structure, the correlation ratio is not applicable for the analysis of time series.

This naturally brings up the question of the elimination of the other factors. Heretofore it has been assumed that the other factors, such as the trend, have been adequately eliminated. In fact such an elimination was necessary to obtain the criteria of steepness and skewness. To compute the values of the criteria, however, it is not necessary to fit an elaborate trend from which to obtain the deviations. All that is needed is the removal of the secular trend and other cyclic movements, except that under investigation, so that the *shape* of the cycle can be determined. Frisch[19] suggested, and Maverick[20] later developed, a method of successive smoothing lines by a controlled freehand tech-

[19] R. Frisch, "Changing Harmonics and Other General Types of Components in Empirical Series," *Skandinavisk Aktuarietidskrift*, 1928, pp. 234–235.

[20] L. A. Maverick, "Time Series: Their Analysis by Successive Smoothings," ECONOMETRICA, Vol. 1, 1933, pp. 238–246.

nique by which each cyclic component would be isolated. The shorter cycles are eliminated first and then those of higher orders. Such a method is open to criticism and there is usually no a priori reasoning behind such eliminations. But it does supply a quick and easy method for the determination of the classification criteria although it is possible that the criteria may be influenced by the personal element in such a freehand method. Unless the series is complicated, however, by numerous cycles of various periods, this method is adequate for the purpose. The method is used for determinining the *value* of the criteria only; it is not used in the final regression functions. Frisch and Waugh[21] have shown that it is not necessary actually to eliminate trend since it can be included as a separate variable in a multiple correlation. In actual practice, therefore, the trends are eliminated as explained by Maverick only for the purpose of determining the values of the classification criteria. After these values have once been determined, they are compared with the values of the other series in the correlation analysis to determine the general form of the regression for each variable, including the trend. The final regression equation is then computed by the ordinary methods of multiple correlation.

The study of correlation is usually approached through the analysis of frequency distributions, but such an approach does not seem applicable to the analysis of time series. Leavens[22] has shown the form of the frequency distribution of smooth periodic curves. These are far from normal distributions except for certain types rarely encountered in statistical series. A sinusoidal series will have a U-shaped distribution and actual time series usually show this tendency in a very flat, or bimodal distribution. Also the analysis of frequency distributions assumes independence of observations and this assumption is not fulfilled in time series. Any single observation chosen at random may meet such requirements but in ordinary analysis at least one item of a time series is known and the probability of the value of the following item is not distributed in accordance with the general frequency distribution but in a much more narrow range. A study of the general distribution of time series shows that the relative steepness of the cycles may be computed but that the computation of the skewness is very complicated and in either case these computations are much more difficult than the ones given in Section II. Therefore, it does not seem that correlation analysis of time series can be approached adequately through a study of their frequency distributions alone.

[21] R. Frisch and F. V. Waugh, "Partial Time Regressions as Compared with Individual Trends," ECONOMETRICA, Vol. 1, 1933, pp. 387–401.

[22] D. H. Leavens, "Frequency Distributions Corresponding to Time Series," *Journal of the American Statistical Association*, Vol. 26, 1931, pp. 407–415.

V. SUMMARY

To summarize very briefly, it has been shown that time series can be classified in two general classifications, one of steepness and one of skewness, and that the general form of the regression function will depend on the comparison of these classification criteria. In general, the regression will be nonlinear, and if the resultant skewness is large there will be a hysteresis loop in the regression function which can be corrected only by the use of two functions, one during the upward trend of the cycle and one during the downward trend. By tracing the pattern formed in the scatter diagram by connecting successive points, the limits within which the optimum lag lies can be determined quickly by noting the direction of rotation of this pattern; for positive correlation, clockwise rotation indicates that the Y variable precedes the X variable, and counterclockwise rotation, the opposite. This rule is reversed in the case of negative correlation.

It must be borne in mind that the coefficients of curvature and skew hysteresis defined above can give only the general form or shape of the regression function and not its analytical type. The actual choice of the regression function is still undetermined. But statistics is a practical study and though we cannot deny the existence of nonlinear regression, we must not run riot with mathematical formulae. The above analysis gives an a priori basis for the general shape of the regression function, but there is no such basis for choosing any of the many functions that would approximate this form. Therefore, until a more adequate foundation of time-series analysis is developed, simple regression functions should be chosen and usually second- or third-degree polynomials or logarithmic functions will suffice.

The above analysis is, in many ways, preliminary. There is an obvious need for the sampling errors of the classification criteria and the coefficients of curvature and skew hysteresis. In fact, the discovery of these sampling errors might easily change the choice of the criteria. But the above analysis does show the reasons underlying the problem of nonlinear regression in the correlation of time series, and it is hoped that further research will put the method suggested here on a sound statistical basis.

HERBERT E. JONES

*Cowles Commission for
Research in Economics
Colorado Springs*

NOTE ON THE PHASE DIAGRAM OF TWO VARIATES

By Ragnar Frisch

When two variates x_1, x_2 move in time in a wavelike manner maintaining approximately the same phase relation (in which case they must also have approximately the same length of swings), the historical path in the (x_1, x_2) diagram will, as is well known, have more or less the shape of an ellipse. This ellipse—the *phase diagram* of the two variates—is useful for various purposes in connection with the study of the time shape of the variates. Mr. Herbert E. Jones in his paper in this issue of Econometrica,[1] has studied some of these uses. There are one or two other points in this connection which may deserve mention.

As a conceptual pattern consider two sine curves with the same frequency α, but in general different phases and amplitudes—the latter may be not even commensurate:

$$x_1 = A_1 \sin(a_1 + \alpha t), \qquad x_2 = A_2 \sin(a_2 + \alpha t), \tag{1}$$

where $a_1 - a_2$ is the *lead* or *lag* of x_1 over x_2 according as this difference is positive or negative. The *lead fraction*, h, is the lead expressed as a fraction of the common period:

$$h = \frac{a_1 - a_2}{2\pi} = \frac{\text{lead}}{\text{period}}. \tag{2}$$

Since the numbering of the two variates is arbitrary we may let x_1 be the leader and consequently take h as a number between 0 and $+1$.

The historical (x_1, x_2) path will now be an exact ellipse as in Figure 1. In an actual case we shall have wiggles around the underlying ellipse.

There are two characteristics of the ellipse which it is fairly easy to determine by a freehand smoothing or simply by a visual inspection, namely, its *thickness* as determined by the ratio $\lambda = c/C$ of one of its axes (for instance the minor axis) to the other (the major axis), and its *inclination* as determined by the slope $\nu = A_2/A_1$ of the diagonal of the circumscribed rectangle. These graphically determined characteristics reveal amongst others the lead or lag between the two time curves. We have indeed

$$\sin 2\pi h = \frac{\nu + \dfrac{1}{\nu}}{\lambda + \dfrac{1}{\lambda}}. \tag{3}$$

[1] Pp. 305–325.

This expression is unchanged if ν is replaced by $1/\nu$, and it is also unchanged if λ is replaced by $1/\lambda$. Hence it does not depend on the numbering of the variates nor on which of the two axes is expressed in terms of the other. Conventionally the numbers λ and ν should here be taken positive in accordance with the convention that the lead h is taken as a number between 0 and 1. For many years I have used (3) as a convenient and rapid means of determining average lead or lag.

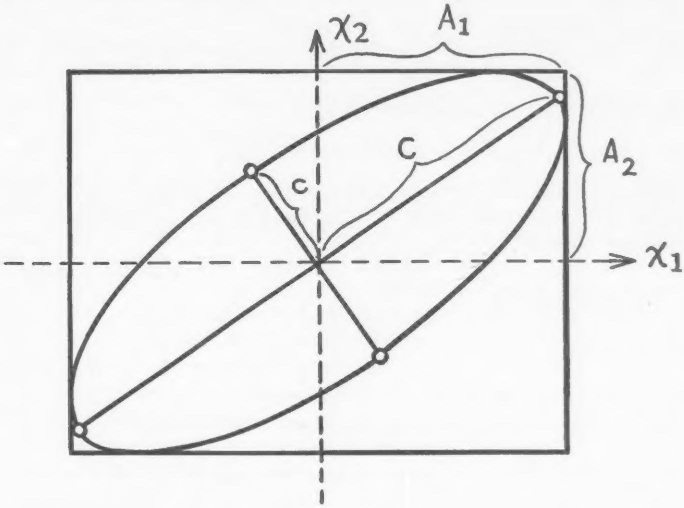


FIGURE 1.

It is easy to prove (3) by noticing that the equation obtained by eliminating t between the two equations (1) is

$$\left(\frac{x_1}{A_1}\right)^2 - 2\frac{x_1x_2}{A_1A_2}\cos 2\pi h + \left(\frac{x_2}{A_2}\right)^2 = \sin^2 2\pi h. \tag{4}$$

Since the quadratic form in the left member is positive definite, (4) represents an ellipse. The square of the ratio between its two axes is

$$\lambda^2 = \frac{A_1^2 + A_2^2 - \sqrt{(A_1^2 + A_2^2)^2 - 4A_1^2A_2^2\sin^2 2\pi h}}{A_1^2 + A_2^2 + \sqrt{(A_1^2 + A_2^2)^2 - 4A_1^2A_2^2\sin^2 2\pi h}}. \tag{5}$$

By the definition of ν we further have

$$\nu + \frac{1}{\nu} = \frac{A_1^2 + A_2^2}{A_1 A_2}. \tag{6}$$

This gives (3).

Instead of determining the ellipse by a freehand smoothing, we may, if a sufficiently large range of observation is available, do it by computing the standard deviations σ_1 and σ_2 and the correlation coefficient r_{12} of x_1 and x_2 over the range. If the range is large as compared with the length of the period, the equation of the ellipse will be

$$\left(\frac{x_1}{\sigma_2}\right)^2 - 2\frac{x_1 x_2}{\sigma_1 \sigma_2} r_{12} + \left(\frac{x_2}{\sigma_2}\right)^2 = 2(1 - r_{12}^2). \tag{7}$$

Hence, the lead will be determined by

$$r_{12} = \cos 2\pi h. \tag{8}$$

This latter formula, however. assumes that the range is fairly large, while (3) may be used whenever a freehand or other smoothing of the ellipse is possible.

RAGNAR FRISCH

University of Oslo

AN EXAMPLE OF MEANINGFUL CURVILINEAR REGRESSIONS IN ECONOMIC TIME SERIES[1]

By Horst Mendershausen

In a recent enquiry into the influence of the business cycle and of weather upon the size of seasonal fluctuations in the building industry[2] the problem arose, as to what shape should be chosen for the regressions connecting the variables involved. This problem is fundamental to all correlation analysis.

The linearity of the regressions, which is assumed frequently, though not always very consciously, will, upon closer examination, be found not rarely to give a bad representation of the relations to be determined. Furthermore, there is often not sufficient theoretical reason for assuming a linear connection.

When, on the other hand, the attempt is made to fit curvilinear regressions to the empirical scatter, the same two difficulties not infrequently arise: (1) even the curvilinear regression, especially when a simple mathematical curve is taken (e.g., a third-degree parabola, a logarithmic curve, or a simple hyperbola), does not fit the scatter; (2) even if the curvilinear regression, whether it is an analytical function or not, does fit the scatter, the result remains unsatisfactory as long as the particular regression shape cannot be explained in terms of the nature of the phenomenon studied.

In what follows, a case is presented in which it has been possible to determine curvilinear regressions which not only fit the scatter well but also possess a clear significance in terms of the phenomenon in hand.

The material of the enquiry consists of the following monthly series for unemployment in the building industry: United Kingdom, 1919–1935; the Netherlands, 1918–1935; Belgium, 1920–1935; Denmark, 1918–1934, and Germany, 1919–1933. The percentages unemployed were transformed into percentages employed, and from these series 12-months moving averages were computed. Gross seasonal indices were then obtained by dividing each value of the original series (the percentage employed) by the value of the 12-months moving average centred for the same month. These ratios were taken as the dependent

[1] The author is much indebted to Dr. Hans Staehle for his kindness in translating this paper into English and for his many valuable suggestions.

The paper was read at the European Meeting of the Econometric Society in Oxford, September, 1936. The author would like to express his thanks for the suggestions which were put forward in the discussion, particularly those of Prof. Ragnar Frisch and Dr. Jakob Marschak.

[2] This enquiry forms the object of a doctor's thesis at the Geneva University which is soon to appear under the title *Les variations de la saisonnalité dans l'Industrie de la construction.*

variable (X_1). Different independent variables were taken for the months in different seasons. For the winter months, January, February, and December, two independent variables were generally considered: (1) the average level of employment during the preceding months of September and October, reflecting mainly the cyclical position of employment at the end of the good season, the seasonal forces then being in an approximate equilibrium (X_2); (2) the average temperature during the particular month (X_3). Different independent variables were used for the summer months from May to August and for the remaining months. See below.

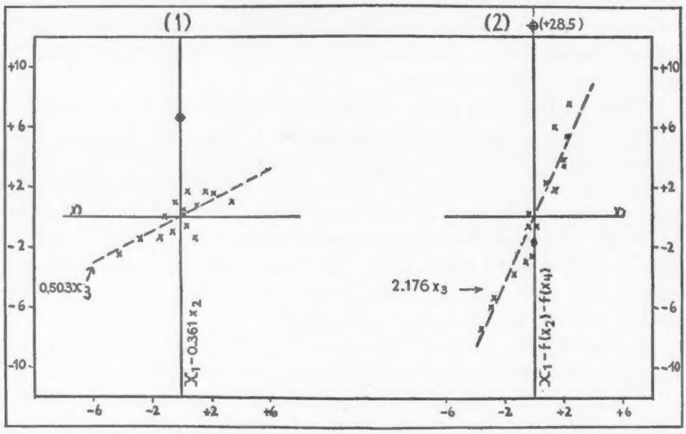


FIGURE 1.—Examples of Temperature Regressions. (1) United Kingdom, 15 January values; (2) Netherlands, 17 January values.

Multiple correlations were computed for each winter month, first on the assumption of linear partial regressions of X_1 on X_2 and on X_3 so that the expected value of X_1, i.e., X_1', for the month of January was determined by

$$X_1' = \overline{X}_1 + b_{12.3}\, x_2 + b_{13.2}\, x_3$$

where x_2 and x_3 are deviations of the values of X_2 and X_3 from their respective arithmetic means, $\overline{X}_1$ is the mean of the values of X_1, and $b_{12.3}$ and $b_{13.2}$ are the partial regression coefficients.

The partial regressions of X_1 on X_3 were well represented by straight lines.

The coefficients $b_{13.2}$ all had positive signs. This corresponded to expectations regarding the connection between the seasonal indices and temperature during the winter months: warm temperatures give rise

to relatively small seasonal movements, i.e., high gross seasonal indices, whereas cold temperatures give rise to relatively large seasonal movements, i.e., low gross seasonal indices. Since there were no a priori reasons for preferring, over the range of the data, any other shape for the regression to the straight line (see also footnote 8 below), the latter was taken.

The case was different for the partial regressions of X_1 on X_2. Here there were a priori objections against the use of straight lines.

Let us first consider the possible range of variation of X_1 and X_2. Whereas temperature undergoes variations to which there is no limit theoretically, the variables X_1, X_2, possess only a limited range of possible variation. X_2, describing the cyclical situation, cannot be larger than 100 per cent (meaning full employment during the months of September and October), and it cannot be smaller than 0 (complete unemployment in these months). We therefore have

$$0 \leqq X_2 \leqq 100.$$

As regards X_1, the seasonal index for a winter month, it cannot fall below 0 per cent of the corresponding 12-months moving average, and it also cannot, at least practically, be larger than 100 per cent. While indeed the winter minimum of employment in the building industry might perhaps, though not easily, be eliminated in our latitudes, it could not rationally be changed into a winter maximum. We therefore practically have

$$0 \leqq X_1 \leqq 100.$$

If there exists a correlation between X_1 and X_2 the line of regression of X_1 on X_2 cannot be either perpendicular or horizontal and must therefore be inclined in some direction and thus approach one of the four corners of the field of variation. These four corners are:

$$\begin{aligned}
&(1)\ X_2 = 0,\ X_1 = 0,\\
&(2)\ X_2 = 0,\ X_1 = 100,\\
&(3)\ X_2 = 100,\ X_1 = 0,\\
&(4)\ X_2 = 100,\ X_1 = 100.
\end{aligned}$$

It must, however, be noted that there is no inherent (i.e., technically determined) reason why any of these limiting values should necessarily be attained. Further hypotheses concerning the shape of the regression curve can be made. Let us consider the total building activity to be composed of two kinds of work: the urgent repair and maintenance work[3] (u) on the one hand, and the postponable work (p), including

[3] That is, work whose postponement would render existing structures uninhabitable or otherwise useless, and work necessary for the mere maintenance of a building enterprise.

expansion work and postponable repair work, on the other. Then each gross seasonal index of employment (X_1) can be taken as a weighted average of two gross seasonal indices, namely the gross seasonal index of urgent work (w_u/ϕ_u) and the gross seasonal index of postponable work (w_p/ϕ_p), i.e.,

$$X_1 = \frac{w_u + w_p}{\phi_u + \phi_p},$$

where the w's represent the level of employment applying to each kind of work and the ϕ's the respective levels of the 12-months moving averages, which serve also as weights.

The value of X_1 thus depends on the values and weights of two partial seasonal indices. We can make some assumptions with regard to these factors, by taking into consideration the known conditions of the building trade.[4]

(1) The urgent work can—roughly speaking—be considered as independent of seasonal changes. This kind of work must be done whatever the climatic conditions are. Thus w_u/ϕ_u may be taken constantly as approximately 100 per cent,[5] i.e., indicating an absence of seasonal fluctuations of employment. Cyclical changes have practically no influence on this partial seasonal index.

(2) The postponable work is influenced by seasonal changes, but the degree of seasonality depends on the cyclical situation, the natural influences being eliminated and other social influences being assumed negligible in a first approximation. For a winter month w_p/ϕ_p can fluctuate between 100 and 0 per cent. We may expect only a *slight* seasonal setback (i.e., high gross seasonal indices) in the winter during a *favourable* period of the business cycle, because high profits and pressing demand make it worth while to undertake building even at the higher costs entailed by work during the winter. If the cyclical situation becomes maximal ($X_2 = 100$), w_p/ϕ_p will become 100 per cent, provided that there is no irreducible minimum of seasonality (due to habits of builders or consumers, lack of technical equipment, etc.). On the other hand a continued depression will call forth a continued concentration of the postponable work in the good season, when building

[4] See S. Kuznets, *Seasonal Variations in Industry and Trade*, New York, 1933; E. Gjermoe, "The Seasonal Movements of Employment in their Relation to Business Cycle," *Nordic Statistical Journal*, Vol. 3, 1931, p. 532 ff., and my own study soon to appear in the *International Labour Review* on the reasons for the undiminished persistence of seasonal variations in the building industry.

[5] Small deviations from this value may occur, not only as a result of random fluctuations but also through changes in the rate of growth of the total number of buildings in existence. These deviations may be neglected in this connection.

costs are lower and risks smaller. Thus w_p/ϕ_p (for a winter month) will fall, and there appears to be no reason for this fall not to reach the value $w_p/\phi_p=0$ during a very bad cyclical situation.

(3) From considerations (1) and (2) one might expect that X_1 would rise or fall according to a rise or fall in X_2, if the weight of the two components were constant. In this case indeed the presumed positive regression of w_p/ϕ_p on X_2 would cause a positive regression of X_1 on X_2 over the main part of the X_2-scale. With falling X_2 this regression would however approach, and stay at, a minimum value

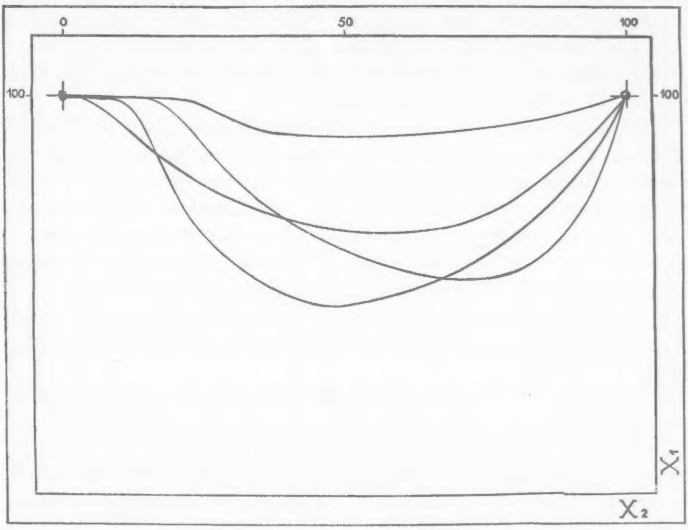


FIGURE 2.—Plausible Regressions of X_1 on X_2 and their limiting values.

$w_u/(\phi_u+\phi_p)$, until $(\phi_u+\phi_p)$ becomes 0, and X_1 therefore an indeterminate value $(0/(0+0))$.

But the weights of the two components of X_1 cannot be assumed to be constant. Cyclical changes have a very small effect on the amount of urgent work but they do affect strongly the amount of postponable work. When building activity increases the postponable work becomes relatively more and more important. Thus the weight of w_p/ϕ_p will increase, the weight of w_u/ϕ_u will diminish relatively, and the regression of X_1 on X_2 will more and more be determined by the positive regression of w_p/ϕ_p on X_2. Under the very probable condition that from a certain moment during the improvement of "times" the increase of

the expression $w_p/(\phi_u+\phi_p)$ will become greater than the decrease of $\phi_u/(\phi_u+\phi_p)$, X_1 will rise. Inversely when building activity decreases, the postponable work will become less and less important, and it may be assumed that it will disappear completely if the cyclical index approaches $X_2=0$. Thus the weight of w_u/ϕ_u will increase and the regression of X_1 on X_2 will be more and more determined by the constant: $w_u/\phi_u=100$. As soon as p vanishes, X_1 will become equal to this constant. This will occur before X_2 has reached 0, namely at that point where X_2 is only determined by work of the kind u.

(4) Assuming a continuous and smooth shape, the regression of X_1 on X_2 for a winter month might thus look like one or another of the U-shaped, more or less concave and more or less symmetrical curves shown in Figure 2, approaching two of our four possible limiting values ($X_2=0$, $X_1=100$; $X_2=100$, $X_1=100$).

In the right-hand part of Figure 2 the influence of the positive regression of w_p/ϕ_p on X_2 dominates, in the left-hand part, the diminishing importance of this influence and the increasing weight of w_u/ϕ_u cause a negative regression of X_1 on X_2. Between these two tendencies there must be a more or less extended sphere of non-regression, where X_1 reaches a minimum. The locus and extension of this sphere is determined by the inherent mechanism of the phenomenon in hand.

For the summer months the regression curve should take the shape of an inverted U, as can easily be shown by an analogous demonstration.

Let us now consider the regression curves which have been found empirically.

The analysis of the scatters showed first that for no series do we possess enough points to construct the regression over the whole theoretical range of the values for X_2. The observed ranges of variation were:

for the United Kingdom from 97 to 70 per cent;
for Belgium from 99 to 66 per cent;
for the Netherlands from 98 to 63 per cent;
for Denmark from 96 to 55 per cent; and
for Germany from 99 to 21 per cent.

This leads to the expectation that the German data offer the best chance for finding the true regression, whereas this chance is smaller, and increasingly smaller, for Denmark, the Netherlands, Belgium, and the United Kingdom.

In accordance with these expectations, curvilinear regressions were found for Germany that both fitted the scatter and corresponded to the above hypotheses. Straight lines as well as third-degree parab-

olas proved unsatisfactory both practically and theoretically. Other or more complicated functions were not tried, mainly because the regression outside the observed range of X_2 could not be determined. Without, however, some knowledge of the unknown range in question, there would not be much sense in expressing the regression by a mathematical function. For this reason *semi-mathematical freehand curves* were fitted to the points of reference found by an averaging process analogous to that suggested by Ezekiel.[6]

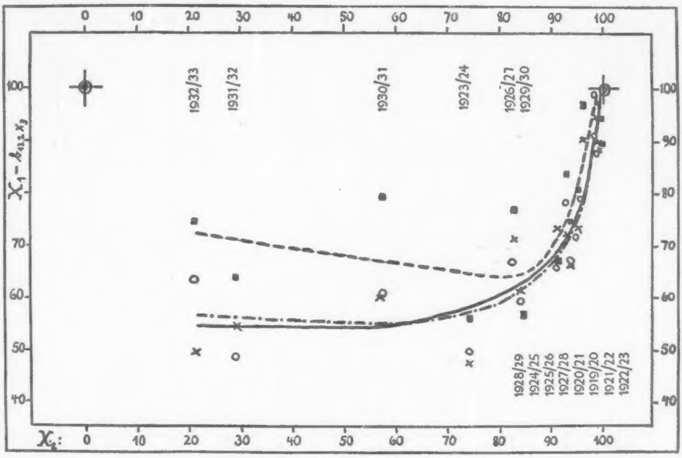


FIGURE 3.—Regressions of X_1 on X_2 for Germany: January, February, and December, 14 values in each case.
January: *o*, —·—·—·—·; February: *x*, ————; December: ■, ------.

The following "coefficients of multiple determination" corrected for the small number of points ($\overline{R}_{1.23}^2$) were found for the various months:

	January	*February*	*December*
Straight line	0.58	0.73	0.38
Third-degree parabola	0.79	0.82	0.51
Semi-mathematical freehand curve	0.88	0.90	0.68

Figure 3 shows the scatter and the regression curves for the three winter months. Whereas the curves are very characteristic in the right-hand part ($X_2 > 80$ per cent) the left-hand part shows a high standard error owing mainly to the small number of points. Figure 4 shows the regression for the month of January together with the

[6] *Methods of Correlation Analysis*, New York, 1930, p. 131.

approximate value of its standard error, computed by the method proposed by Ezekiel.[7]

The systematic way, however, in which the scatter approaches the regression curve, as well as the correspondence between the hypotheses and the real regression, strongly suggest that even the left-hand part of the curve ($X_2 < 80$ per cent) is more significant than would appear from its high standard error. Even if in its details it may be accidental, the left-hand part of the curve nevertheless clearly exhibits a general

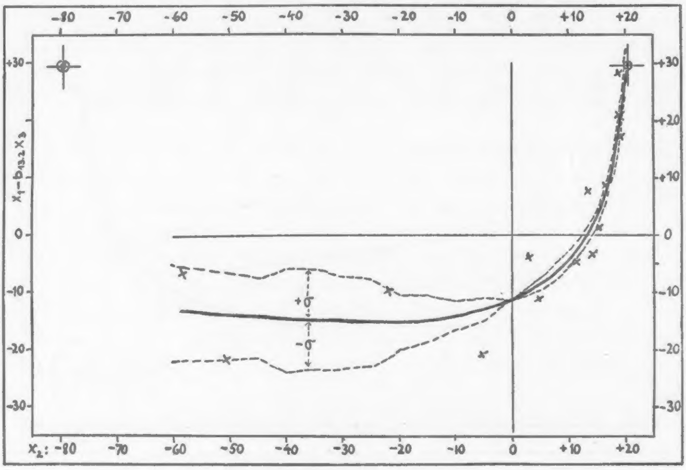


FIGURE 4.—Regression of X_1 on X_2 and approximate values of its standard error for Germany: 14 January values.

tendency for the strong initial drop of the winter indices, which takes place with the increasing cyclical depression, to slow down and finally to stop, and for the indices even to show a tendency toward a renewed increase, rather than toward a continuation of the decline. It may, however, be imagined that if more complete materials were available, the regression might at the lower end of the X_2-scale tend neither toward $X_1 = 0$ nor toward $X_1 = 100$. It might be, for example, that either any intermediate value of the X_1-scale would be found as the real lower limiting value, or that the correlation between X_2 and X_1 completely vanishes in this part of the X_2-scale.

Similar partial regressions of X_1 on X_2 were found for two im-

[7] *Op. cit.*, pp. 263 and 384 ff.

portant professional groups of the building industry, i.e., bricklayers and painters.

The *Danish* material supplied analogous results, with the only difference that, given the smaller range of the cyclical fluctuations, it was not possible to get near the point where $X_2=100$ nor to the other point where the fall of the seasonal indices stops. Figure 5 shows by way of example the scatters and the regression curves for the months of January, February, and December.

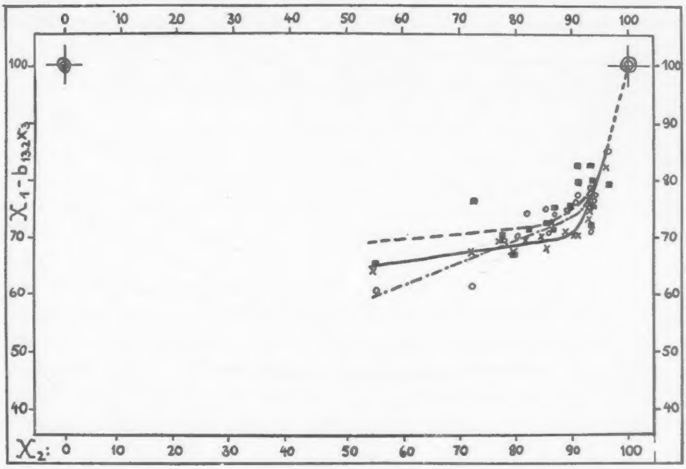


FIGURE 5.—Regression of X_1 on X_2 for Denmark: January, February, and December, 16 values in each case.
January: *x*, ————; February: *o*, –·–·–·–·–·; December: ■, – – – –.

The following table of the values of $\overline{R}_{1.23}^2$ gives evidence of the superiority of the curvilinear regression over the straight line for January and February. It is less striking than in the German case. In December some extreme values are not so well fitted by the curve as by the straight line, so that the correlation becomes even worse.

	January	*February*	*December*
Straight line	0.83	0.83	0.70
Semi-mathematical freehand curve	0.94	0.86	0.60

It has not yet been possible in the analysis of the Dutch, Belgian, and British materials to discover such curvilinear regressions of X_1 on X_2 as would show a better fit than the linear regressions. There are two

reasons for this: (1) the observed range of the X_2 values is smaller; (2) the seasonal indices for the winter months are higher, therefore also the range of the X_1 values is smaller than in Denmark, not to speak of Germany. The narrower the range of the values for X_1 and X_2, the easier it is to obtain a good fit simply by the use of a straight-line regression.[8] Diagram 6 shows the partial regressions for the month of January relating to all the five countries which it has provisionally been possible to cover in this enquiry.

In the case of the United Kingdom the straight line already points toward the upper limit. No tendency toward curvilinearity could be

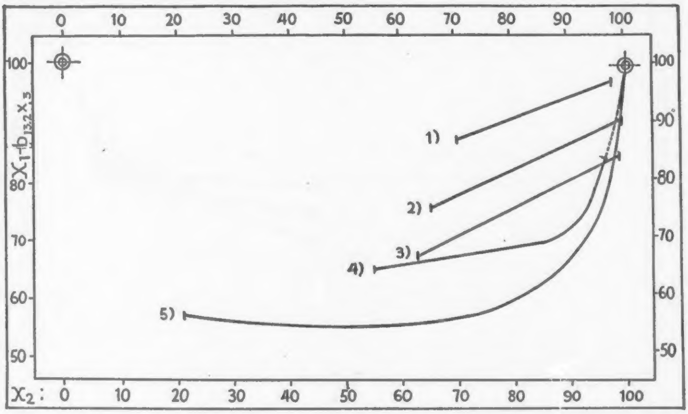


FIGURE 6.—Regressions of X_1 on X_2 in (1) United Kingdom (15 values), (2) Belgium (14 values), (3) the Netherlands (17 values), (4) Denmark (16 values), and (5) Germany (14 values), indicating in each case the range of X_2 covered by the observations.

discovered. This does not exclude the possibility that, if X_2 had smaller values, the fall of the regression toward the left would slow down or even stop and point toward the lower limit (0, 100). In the case of Belgium and the Netherlands there were some traces of curvilinearity, but the fit of such curves was not sufficiently superior to that of the straight lines to permit of the conclusion that the curvilinearity had been proved. It may be assumed that in these cases also this conclusion

[8] This also applies to the partial temperature regressions dealt with above. It is, in fact, very probable that, to the extent to which X_3 attains extraordinarily high or low values so that X_1 will therefore be nearly 0 or 100, the regression $f'(x_3)$ will differ from a straight line. It may be imagined that from the shape / it may change into the shape ∫. This situation however did not occur in our examples. Cf. the position of the points ⊕ ($X_1 = 100$) in Figure 1.

might be arrived at by a refinement of the analysis, the inclusion of further variables, etc.[9] Should this, however, not be the case, a new situation would arise. The regression of X_1 on X_2, since it is seen in our figure not to be pointing in the direction of $X_1=100$, cannot pass through this point. This means that, while the seasonal movement showed a tendency to disappear completely in Germany, Denmark, and the United Kingdom, when X_2 attained its maximum, this would not be the case in Holland and Belgium, although there is a reduction in the seasonal variations corresponding to an improved general economic situation. This question of an irreducible minimum of seasonality could only be cleared up by a further enquiry.

Our hypothesis that the true regression of X_1 on X_2 is curvilinear, with a tendency toward a U-shape, corresponds to the relation between the average seasonal deviation for a given year and the central value of the 12-months moving average of the same year, which has been found by E. Gjermoe.[10] Gjermoe has worked on the data for the percentage employed among: (1) all trade-union members; (2) metal workers; and (3) "seasonal" (i.e., mainly building) workers, of Norway.

Gjermoe fitted hyperbolas to his scatters. Obviously, however, such hyperbolas are inconsistent both with the existence of definite finite limits for the extreme values of the variables and also with a U-shaped regression. Both these conditions should also have been taken into account in his enquiry, for in fact the scatter shown in Gjermoe's diagrams 1 and 3 shows a clear tendency to deviate from the hyperbola and to approximate an inverted U-shape curve. (The inverted "U" follows from the use of the average seasonal deviation for the whole year instead of seasonal indices for particular winter months as the independent variable). In his diagrams 1, 2, and 3, furthermore, a clear tendency of the points toward the upper limiting value can be seen.

Gjermoe's most interesting contribution may therefore be criticised in this connection on the ground that the hyperbolic regression was assumed without any consideration of the conditions of the phenomenon studied. Furthermore, the choice of the seasonal deviation for entire years, the omission of the forces of nature influencing employment, and a few other points might be criticised.

The presence of curvilinearity with a tendency toward a U-shape in the Norwegian series studied by Gjermoe not only offers a further example of the regression shape discussed in this paper, but it also

[9] It may be mentioned that it has been possible to improve the Dutch correlation quite considerably by including the quantity of rainfall as a third independent variable (X_4).

[10] *Op. cit.*, pp. 544 ff.

indicates that this shape might possibly be met with in industries other than the building trade.

As regards the summer months it has been possible to find regressions for Germany that are exactly analogous to those found for the winter months. They are shown in Figure 7 for the months of June and July.

In these computations only one single independent variable (X_5) was used, i.e., the percentage of employment for April, reflecting mainly the cyclical situation at the beginning of the good season. (In

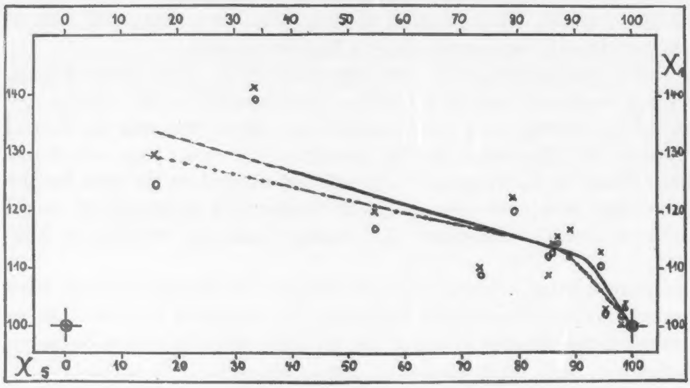


FIGURE 7.—Regression of X_1 on X_5 for June and July in Germany (13 values in each case).
June: o, – · – · – ·; July: x, ————.

most of the series the gross seasonal index of April varied very little and proved to be uncorrelated with the cyclical and seasonal situation during the winter. Therefore, the percentage of employment for this month could in these cases be taken as an approximate index of the cyclical situation at the beginning of the good season.)

The situation is quite analogous to that concerning the winter months. X_1 may vary from 100 to ∞, but rather shows a tendency to remain near the limiting value $X_5=0$, $X_1=100$, when X_5 becomes relatively small. The available points permit of drawing the beginning of the curve which might, if it were possible to extend it over the whole scale, conceivably take an inverted U-shape. The following table shows the slight improvement of the correlation by this curve:

	June	*July*
Straight line	0.64	0.68
Semi-mathematical freehand curve	0.67	0.72

The left-hand part of the curve ($X_5 < 80$ per cent) is again uncertain because of the lack of observations. We applied finally Blakeman's test of linearity to the analyzed regressions. The curvilinearity was found to be "significant" in the German winter cases, "insignificant" in the Danish winter and the German summer cases. This test is, however, too rigorous to be taken as decisive in this connection. The stated curvilinear regressions correspond better than the linear to our expectations, and there is no one case in compensation where a curvilinear tendency in another than the supposed sense could be secured.

In conclusion, it may be said that the material proves that a curvilinear regression of X_1 on X_2, for which plausible hypotheses can be made, exists in Germany and in Denmark. The regression could not be determined over the very lowest part of the X_2-scale. The German materials for the winter months, however, make our hypothesis of a U-shape for the total regression probable.

Horst Mendershausen

Geneva

ECONOMIC CONSEQUENCES OF TECHNICAL DEVELOPMENT, WITH SOME ILLUSTRATIONS FROM DANISH INDUSTRIES

By L. R. Nienstaedt

1. COSTS AND THE GENERAL CONCEPTION OF TECHNICAL DEVELOPMENT

In a recent publication by Schneider[1] the problem of technical development and its relation to and influence on capital investment is taken up for consideration. Schneider makes reference to studies by Gustaf Åkerman[2] and from this investigator he has acquired the term "automatization," which is used as a designation of the essential features of technical development as far as production is concerned. For the sake of clearness the definition of the term as offered by Åkerman is modified by Schneider. He defines[3] automatization by saying that "a mechanical work process has attained a higher degree of automatization with regard to the capital invested in machinery, when the production of equal quantities of product in a certain length of time (e.g., an hour) is accomplished with a smaller number of workers cooperating now with the new and more complicated machines."

On the basis of an analysis of total costs in relation to various causes of depreciation or deterioration as well as in relation to capital invested, interest paid on the investment, and total yearly wages (depending on total hours of utilization) he formulates a system of equations corresponding to entrepreneurial calculations made, when in practical life a decision as to the advisability of an investment is needed.

We shall attack the problem differently, giving it a more comprehensive formulation, which is not confined to the individual production unit at a certain time, but which will explain to us why technical improvements, so to speak, force themselves upon an industrial community and step by step change the conditions of this community in the manner observed in all modern societies. For it should well be remembered that once a certain technical possibility has been put at the disposal of a certain industry the question of utilizing it in a certain production unit is not a *choice* but a *necessity* and the moment at which the step is made may determine once and for all if the particular unit in question shall continue to operate or not.

The equations as formulated by Schneider do not allow the con-

[1] "Das Zeitmoment in der Theorie der Produktion," *Jahrbücher für Nationalökonomie und Statistik*, Jan., 1936, pp. 45 ff.

[2] G. Åkerman: *Realkapital und Kapital*, Hft. 1., Stockholm, 1923.

[3] *Op. cit.*, p. 54.

sideration of this aspect of technical development nor do they allow the consideration of relations between strictly technically determined quantities and money values or economic quantities; but a slight change in the basic definitions and formulation of the problem may accomplish the end in view.

2. THE WORK PROCESS AND TECHNICAL DEVELOPMENT

It appears that the term "automatization" in the very general sense in which it is taken by Åkerman and Schneider does not cover the conditions of practical life, for it has reference only to a certain *step* in the development of *mechanical* processes (installation of automatic machines) but is not at all significant to technical development in the early stages. For this reason Åkerman has to distinguish between two phases of which the first one is evidently not an "automatization" as far as the word is used technically. Further the word has no significance in relation to the development of *nonmechanical production processes* such as, e.g., the smelting of pig iron, and, as we shall have an opportunity to see, the essential facts of technical development apply in quite the same manner to the latter processes as to the former. For this reason, and as a first generalization, we shall prefer to describe technical development, as far as production is concerned, simply as increasing productivity, i.e., *increasing quantity of product per productive man-hour*, and it will be the purpose of the present analysis to distinguish between various ways in which this increase can be accomplished by technical means.

The performance of a physical work process is the basis of all production. By a work process we will understand any kind of transformation or transportation of material substance, which has an economic end in view and by which physical energy is consumed or liberated. In the first place we will distinguish between a *primitive work process* and an *entirely technical work process.* In *a primitive work process* the necessary work energy is supplied by man himself or animals. It is always mechanical. In an *entirely technical work process* the necessary work energy is supplied by technical means only, e.g., heat. It may be either mechanical, since heat energy by transformation in mechanical power generators (steam engine, turbine, diesel engine, etc.) may be transformed into movement, or it may be nonmechanical in a physical sense, when heat is directly applied to the work process. It may be nonmechanical in a chemical sense, necessary energy for the process being supplied: (1) by the reacting materials themselves with no excess or deficiency of energy; (2) by energy to be applied from without (endothermic processes); (3) by energy developing in excess (exothermic processes).

Mechanical energy supplied by technical means may co-operate directly with mechanical energy supplied by man as an integrating part of a certain work process. The work process is then nonautomatic. This is a *combined technical process* and corresponds to the first phase of "automatization" as described by Åkerman. It replaces a primitive work process.[4] By removing the direct co-operation of man a combined technical process develops into an *entirely technical process*. It is then fully automatic. All work energy for the mechanical process itself is then supplied by technical means. This corresponds to "automatization" and covers the current technical interpretation of the word.

3. THE QUANTITATIVE RELATION OF TECHNICAL AND ECONOMIC QUANTITIES EXPRESSED AS AN EQUATION

To form a conception of the production process and of its relation to economic quantities we may consider various stages of production numbered, say, 1, 2, 3, 4. At each stage value-product is added. At the final stage the material has been worked into the highest state of finish and is sold for consumption and put to use. When buying the goods the consumers spend a certain amount of money equal to the total value of the product. It is the counterpart of all costs and profits realized at every stage, which the material has passed from its point of origin in nature to consumption.

At each stage work energy must be spent. This energy may be either human or technical or both. *Neither material nor money values will pass any stage without an expenditure of work energy* and all movements are in one and the same direction always, money values and materials moving opposite to one another, however.

The quantities in question may be measured: Money values of course in monetary units, material in weight units or pieces, work energy in man-hours, man-weeks, or man-years if it is human, and in corresponding units if it is technical energy (e.g., kwh., kg-cal. = kg.°).

The price level p of any stage is the relation of *total value of product*, v, leaving the stage to the *total quantity of product*, x; i.e., $p=v/x$.

The productivity, μ, of any stage is the relation of the total quantity of product, x, passing the stage to the total productive man-hours, b, necessary to make the product pass; i.e., $\mu=x/b$.

When the manufacturing technique of a certain process is improved the aim is to increase the productivity. This is accomplished by investing capital in apparatus and machinery transmitting technical energy.

To account for such a development in a technical-economic sense we

[4] Evidently the work process in this case must be wholly mechanical, for man can only co-operate as an integrating part of a *mechanical* work process. All human work in a physical sense is movement.

may use the above simple relations for establishing the necessary equation describing a certain improvement of technique. It may involve a capital outlay of K *per worker* co-operating with the capital. We will consider the simple case when a primitive process develops into a combined technical process. In this case no capital is employed *before* the change takes place. The basis of our equations is the fact that total income must cover all outlay, disregarding profits. If the productivity of one man in one hour is μ_1, and the price of the finished product P, then total income per man in one hour is $\mu_1 p$. The costs are his wages, L, plus outlay on raw materials, $m_1 p_1$ (m_1 = quantity used per man-hour, p_1 its price). The equation for the primitive technique will be

$$\mu_1 P = L + m_1 p_1.$$

The quantities refer to one hour. If we wish to refer them to one year, and h hours are worked per man in a year, then the equation is

$$\mu_1 Ph = Lh + m_1 p_1 \cdot h. \tag{1}$$

If K is the capital invested *per man* in the new technique, which corresponds to the productivity μ_2, we have to consider the costs of investment besides the outlay on wages and raw materials. The latter will be $L + m_2 p_1$ in one hour, $Lh + m_2 p_1 h$ in a year as before. The costs of investment are: (1) *Interest*, Ki (i = rate of interest) and (2) *depreciation*. The latter has to be considered as a *time depreciation*, taking place irrespective of utilization, over a certain number of years "a." At the end of this period the capital is consumed. In one year the quantity consumed is K/a. Depreciation has to be considered further as *wear due to utilization*. If the total number of hours in which the apparatus in question may be used according to its purpose is H and h the number of hours, during which the apparatus is used in one year, then $K \cdot h/H$ is the total depreciation in one year due to this cause. When all costs, as in the first instance are referred to one year with h hours of work per man, the second equation—under the assumption that prices at the moment remain unchanged—will be

$$\mu_2 Ph = Lh + m_2 p_1 h + Ki + \frac{K}{a} + \frac{Kh}{H}. \tag{2}$$

Deducting (1) from (2) and assuming that the quantity of raw materials necessary will increase in the same ratio as the productivity and that, accordingly, $m_2 = m_1 \mu_2 / \mu_1$, we find

$$h \frac{\mu_2 - \mu_1}{\mu_1} (P\mu_1 - p_1 m_1) = K\left(i + \frac{1}{a}\right) + \frac{Kh}{H}. \tag{3}$$

4. THE FACTORS OF THE EQUATION

This equation seems to contain a number of interesting relations as far as technical development and increasing production are concerned, and to these we shall briefly refer:

$(\mu_2-\mu_1)/\mu_1$ is the relative increase of productivity, in other words a kind of measure of the degree of technical improvement. Åkerman would probably call it the increase in automatization. The dimension, as will be seen, is a pure number. It will be realized that in the case of individual technical improvements it will be of little value to consider a continuous relation $d\mu/\mu$, for all technical improvements individually considered are discontinuous.

$P\mu_1-p_1m_1$ is the value product added *per man and hour* in the primitive technique. The term will read as "money units per man-hour." It is of course equal to total costs per man-hour. If the parenthesis is combined with the divisor, μ_1, the term $(P\mu_1-p_1m_1)/\mu_1$ will read as "the value product added per unit of material." The corresponding term on the basis of the increased productivity, μ_2, would of course be $(P\mu_2-p_1m_2)/\mu_2$. The dimension will read as "dollar per unit of product," as it should, irrespective of the time factor and be equal to the term as found in the primitive technique.

Since the increasing productivity is realized by the investment of a capital, K, it will be of interest to determine the ratio of one to the other. This we may do by dividing both sides of the equation by K:

$$h\,\frac{\mu_2-\mu_1}{K}\quad\frac{P\mu_1-p_1m_1}{\mu_1}=\left(\frac{ai+1}{a}\right)+\frac{h}{H}\cdot \tag{4}$$

In this expression, $(\mu_2-\mu_1)/K$ is the increase of productivity obtained per hour per money unit invested. It will read as "product per hour and money unit" (not per man, for K, capital per man, cancels out against $\mu_2-\mu_1$, product per man and hour).

The dimension of the left-hand side is

$$\text{hours}\,\frac{\text{kilograms/hours}}{\text{dollars}}\quad\frac{\text{dollars}}{\text{kilograms}}=1,$$

which corresponds to the right-hand side, for a and i are pure numbers. There is a point of significance, however, to which it is necessary to call attention. In the relation of dollars to dollars the term above the fraction is not truly of the same nature as the term below, for it will be realized that "dollars" evolving out of the creation of value product are money values "in movement," so to speak, whereas "dollars" invested are "fixed."

5. THE MEASUREMENT OF PRODUCTIVITY AND THE POSSIBILITIES OF TECHNICAL DEVELOPMENT

We have chosen to measure technical development by the increase of productivity, i.e., by the change of the factor $(\mu_2-\mu_1)/\mu_1$ when μ_2 increases. It would be interesting if we were able to do so at various stages of development of the same process of production in the course of its transformation from a primitive to a purely technical work process. Unfortunately we have at present no exact data which will permit this, and we must content ourselves by selecting some examples illustrating characteristic changes and development of production.

Before doing so, we have to consider the general principles of observation and the method to be used for measuring productivity of the different types of work processes.

Regarding the *former* we might of course choose an individual concrete technical change in a *particular* work process as our object of observation, but because of ever-changing divisions of labour it would be difficult to be certain of always considering exactly the same process and to account for exactly that part of total value product added, which this particular work process has contributed to the value of the finished article. In other words the accuracy, which might be gained in this way with regard to details of the technical specification, would be lost when the co-ordination with the economic quantities in the equation we have established is attempted.

Further we have to realize that the "value" of any article at any casual point of manufacture is a very doubtful quantity when expressed in money units, for value and costs are not identical. In a true sense value can only be established in a market, where supply and demand for a product of certain standards of quality interact. The market value may then be above, equal to, or below the costs of production. An unfinished or half-finished product has no market value. This is another reason why we cannot consider the individual work process, in observing the interaction of values and technical changes.

Instead we shall consider a factory, an entire industry, or any clearly definable production stage *as a unit*. In this case all manufacturing costs, all profit, and all expenses *have* already been charged to the product, when it *leaves* the factory or the production stage, and no costs, no profit, and no expenses have been charged as yet, when the materials from which the product is made *arrive*. This principle of observation assumes that technical development in relation to production may be wholly accounted for in terms of man-hours or power-hours. In other words, that the application of energy, human or technical or both, is the essence of all production and that all changes of a technical nature affecting production will affect also the

quantities of energy applied. Only if this is the case is it possible to describe and account for technical changes happening at any point of a whole series of operations necessary to complete a certain product by simply accounting for the man-hours or power-hours going into the work process.

If the method as outlined will serve its purpose it will relieve the economist of a bewildering maze of technical details, and permit a generalization which ought to prove of value on a number of points.

With this in mind we wish then to establish the following principles of observation: All observations of quantities are to be made on a production unit, factory, or industry considered as an inseparable whole. Limits are to be established defining exactly what is "inside" and what is "outside" of the unit of production. Quantities are to be measured as they pass the limits set "into" or "out of" the unit of observation.

Changing quantities are then an indication of changing conditions within the production unit itself or the surrounding community and these changes we propose to explain technically and economically.

When comparing observations of quantities made at different time intervals it is of course necessary that the raw materials or finished product passing the limits of the unit of production in either direction always correspond to the same standard of quality.

We have next to consider the measurement of productivity. In the case of a *primitive work process* the productivity is clearly the relation of total production to total productive man-hours. It can only increase to the point, where the physical power of the human worker is completely spent in the apparatus he uses without any loss due to friction or *idle movements*. This would mean the attainment of 100 per cent efficiency, which is of course an impossibility. Technical improvements will try to advance as much as possible towards this goal, however, but they may consist in improvements of efficiency only. Examples are: The flying shuttle invented by John Kay, and the spinning jenny invented by Hargreaves.

The *entirely technical work process* can be characterized by studying the amount of product obtained per unit of technical energy applied. We shall call it the *yield*. The productivity, μ_p, or amount of product per hour, is then determined by the *amount* of energy applied per hour, which we shall call the *intensity*, i_p, multiplied by the *yield* per unit of energy, y:

$$\mu_p = i_p \cdot y.$$

As in the case of the primitive work process the yield cannot increase beyond the point where every unit of energy applied is transformed into useful work without any loss or at 100 per cent efficiency. The

amount of energy which it is possible to apply in one hour may be increased, however. This procedure we shall call *technical intensification,* and technical intensification may take place as long as it is possible to increase the *speed* of the moving parts performing mechanical operations and the *quantities* of reacting materials in chemical or physical work processes. These quantities, however, are again determined by the possibility of handling them mechanically.

The amount of technical energy it is possible to apply to a certain work process in one hour is dependent on available *capacity* for transmitting energy or the horsepower or kilowatt *effect.* The available capacity is only rarely utilized to the utmost possible limit, i.e., at a full load. The relation of energy, which *has* been transmitted to what *might* have been transmitted at a full load is called *the capacity factor.*

At 100 per cent efficiency the yield of any technical work process is a natural constant determined by the physical or chemical properties of the materials treated. It takes 483 kg-cal. to produce 1 kg. of cement, 2350 kwh. to produce 1 ton of pig iron, 770 kg-cal. to produce 1 kg. of lime from limestone. This natural constant we may call the theoretical yield or *the yield factor, k.* The *yield* of an entirely technical work process is consequently equal to: The theoretical yield k, multiplied by the efficiency η. *The productivity* is equal to: The intensity, i_p, times efficiency, η, times theoretical yield, k:

$$\mu_p = i_p \eta k. \tag{5}$$

It must be understood, that the yield factor is a natural constant only as long as a certain raw material is subject in principle to the same work process producing the same quality of finished product. This is so irrespective of the technical arrangement. When the work process is changed the theoretical yield factor will more often than not remain unchanged as long as the initial and final standards of finish, to which the product and the raw materials from which it is made have to correspond, remain unchanged.

Since the yield factor is a natural constant, technical development must consist either in intensification or in improvements of efficiency or both. Other possibilities do not exist.

In *a combined work process* the human worker is a necessary and integrating part of the technical arrangement. However little his physical power is called upon in finishing a certain operation, he is tied to the machine, which cannot operate without his continued attention. Under these circumstances it is natural to measure productivity, μ_c, as the product per man-hour as in the case of the primitive work process. The influence of the technical energy applied must be taken into consideration, however, as in the case of the entirely technical work process.

Consequently productivity is equal to: Intensity, i_p (in this case, energy applied per *man*-hour), times the yield of the energy:

$$\mu_c = i_c \eta k. \tag{6}$$

As in the case of the entirely technical work process, technical development may consist in intensification as well as in improvements of efficiency. But in this case not only the efficiency of utilization of technical energy, but also the efficiency of the man-hour may be improved as well. This will cause a decrease in the number of man-hours necessary to produce a certain quantity of product and such an improvement will evidently appear as technical intensification, which is wrong. To determine the true cause of the increase of productivity, however, we need only to find the number of man-hours worked in 24 hours before and after the change in productivity. Evidently improving efficiency of the man-hour will cause the total number of man-hours worked in 24 hours to decline, whereas total production and total consumption of technical energy in the same length of time will not change. If they do, it is an indication that changes in the application of technical energy are also taking place to some extent.

Apparently we have thus accomplished the analysis of the various possibilities and principles involved, when the productivity of the man-hour is increased under modern technical conditions, permitting the application of two different sources of energy to the production process. We now shall attempt to apply our method to practical examples.

6. THE STATISTICAL OBSERVATION OF PRODUCTIVITY AND TECHNICAL DEVELOPMENT

Our first example of technical development is an "automatization" in the true sense of the word, i.e., the technical change consists in the exclusion of man labour as an integrating and alternating part of the operations necessary to complete the work. The change is a replacement of hydraulic presses in a vegetable-oil factory by automatic presses. The hydraulic press is charged and emptied by hand labour. The automatic press is worked continuously and is charged and emptied by mechanical means, which convey the material as a continuous stream to and from the machine.

Table 1 contains our observations of technical quantities, whereas money values have to be omitted in this case. Columns (1) and (2) represent "Total production" and "Total man-hours" worked. Their ratio is the productivity (column 3). It will be noticed that, beginning with observation No. 5 (the fifth year) the productivity increased from an average of 100 to an average of 120 in the last three observations, an increase of 20 per cent altogether. This corresponds to the fact, that in

a period covering $1\frac{1}{2}$ years from October 7 to May 16 the oil-pressing department was equipped with modern automatic presses, whilst no particular changes occurred in the immediately preceding and subsequent period.

The automatic presses use more power. This will be evident from column (4), which informs us of the total energy consumption, whilst column (5) is the relation to man-hours or, as we have called it, the technical intensity. It will be seen how the rising productivity corresponds to a considerable increase in this ratio, i.e., the technical change must be characterized as intensification.

TABLE 1
INTENSIFICATION

Observation (Year)	Total Production	Total Man-Hours	Productivity	Total Energy Consumption	Intensity	Yield	
No.	kg. (1)	Mh. (2)	kg./Mh. (3)	kg.° (4)	kg.°/Mh. (5)	kg.° (6)	
1	100	100	100	100	100	100	
2	133	128	104	116	91	115	
3	99	107	93	100	94	99 Oct. 7	
4	87	85	102	84	99	104	Automatization
5	91	81	112	92	114	99 May 16	
6	99	84	118	102	121	97	
7	108	90	120	112	124	97	
8	112	91	123	112	123	100	

If we wish to interpret the rising productivity in terms of technical energy according to our previous analysis, we must write it as the product of intensity and the yield. In column (6) the latter has been determined from columns (1) and (4). It will be seen how the efficiency of application declines and with it the yield as the intensity increases, but of course in a different ratio and not very much. Before the change of technical conditions it is 105 on an average and after the change it is 98. No doubt such an occurrence will be quite common when high-speed machinery is introduced. We are unable in the present case to determine a theoretical yield and in most cases it will be difficult to do so for mechanical work processes, but it should be possible, for the efficiency of transmission for any piece of machinery is always determinable.

Having thus far illustrated an instance of technical development,

which is typical of automatization, we will consider also a typical case showing improvements in the efficiency of application of technical energy. This we may do in considering a different department of the same vegetable-oil factory used in the first example.

In this case the technical process consists in the extraction of soya beans by means of benzine. This process involves the application of benzine to soya beans in large vessels known as extractors and the subsequent fractured distillation of the solution of soya bean oil in benzine separating thus the oil from the benzine. Apart from relatively small amounts of energy supplied to the rolling mills and the stirring apparatus the greater part of the energy necessary in the production process is supplied as heat in steam. This makes it possible to calculate the yield factor and efficiency of the process very closely.

TABLE 2
DETERMINATION OF YIELD FACTOR AND EFFICIENCY

Operation No.	Theoretical Energy Consumption	Actual Energy Consumption (measured)	Efficiency
	10^6 kg-cal.	10^6 kg-cal.	
1	4.50	5.07	0.89
2	5.44	7.24	0.75
3	24.00	36.20	0.66
4	23.00	94.80	0.24
5	15.00	18.80	0.80
6	6.00	7.24	0.83
Total	77.94	169.40	0.46

In Table 2 will be found the calculation of energy *theoretically* necessary to bring 400 tons of soya beans through 6 operations and produce raw oil and meal. At each operation is further put down the *actual* consumption of heat energy as determined from measurements of steam consumption in the various processes made at the factory. The ratio of the first of the two figures to the second is then the efficiency of the process. It will be seen how widely the efficiency of the various operations vary. In some cases it is not far from the absolute maximum, showing 89 per cent efficiency. When the theoretical heat quantity needed to produce 400 tons is 77.94×10^6 kg-cal. or 80 million, the yield factor is evidently 5 tons per million kg-cal.

We may then proceed to make observations of the production unit in question. Table 3 contains the figures in index numbers and covers a period of 6 years. Both the productivity and the intensity rise. In the beginning there is a slight decline in a year when total production also

decreases, but in the following year they rise quite definitely. The efficiency declines from year to year until the fourth year, when it rises suddenly. This should indicate a definite change in technical conditions, which actually it does, for in the period from June 29 to September 26 a new process of distillation is introduced, which represents a considerable saving in steam, i.e., our indicators have revealed an improvement in the efficiency of application of technical energy as was to be expected from the nature of the technical change.

At the same time it is evident that it is not the only cause of rising productivity, for intensity continues to increase. This, however, does likewise correspond to actual facts for in an endeavour to press the capacity of the installation as much as possible so as to keep up with rising demands from the sales department, steam consumption has

TABLE 3

IMPROVEMENT OF TECHNICAL EFFICIENCY

Observation (Year)	Total Production	Total Man-Hours	Productivity	Total Energy Consumption	Intensity	Yield	
No.	kg. (1)	Mh. (2)	kg./Mh. (3)	kg.° (4)	kg.°/Mh. (5)	kg./kg.° (6)	
1	100	100	100	100	100	100	
2	83	87	96	82	94	101	
3	106	101	105	110	109	96	
4	114	103	111	95	92	120	June 29
5	111	92	121	101	110	110	Sept. 26
6	128	97	131	155	160	83	

risen considerably at another point of the process. Raw materials going through the work process are so preponderantly soya beans (90–95 per cent) during the whole period of observation, that no changes of the physical conditions of the raw materials have occurred.

7. THE STATISTICAL OBSERVATION OF VALUE PRODUCT ADDED AND THE ADJUSTMENT OF QUANTITATIVE RELATIONS

To get a more complete understanding of the interaction between the various factors we shall reconsider equation (4). We put

$$\frac{\mu_2 - \mu_1}{K} = q', \qquad \frac{P\mu_1 - p_1 m_1}{\mu_2} = \alpha,$$

which will change (4) into

$$q'\alpha h = \frac{ai+1}{a} + \frac{h}{H},$$

$$h\left(\alpha q' - \frac{1}{H}\right) = \frac{ai+1}{a}. \tag{7}$$

Since $(ai+1)/a > 1$ we have $\alpha q' > 1/H$, which expresses the natural condition, that additional income obtained in one hour by investing the sum K per worker must at all events be larger than the deterioration of the capital through utilization in the same length of time, a condition which is of course always fulfilled.

If the origin of the system of co-ordinates is changed to $q' = 1/\alpha H$, the equation will read

$$\alpha hq' = \frac{ai+1}{a}. \tag{8}$$

As far as a and i are constants, this equation represents an equilateral hyperboloid. It expresses the fact that a difference in the price levels of raw materials and finished products, permitting value product added per unit to be α with a productivity q' in h hours, will produce an income sufficient to cover necessary expenses of investment and depreciation. Of course this is a lower limit and the aim will always be to obtain as much income above this limit as possible, just as the possibility exists that the income attained is insufficient to satisfy the condition of the right-hand side. Consequently the equation should properly be written as an inequality as follows:

$$\alpha hq' \gtreqless \frac{ai+1}{a}.$$

Now the value product added in manufacture will have a tendency to decrease in the course of time. With unchanging technique and limited possibilities of utilization of capacity this will endanger the profit or income realized above the limit and bring the relation of quantities on the left-hand side nearer to the point of equilibrium expressed by the sign of equation. Long before this happens, however, an attempt will be made to counteract the influence of a declining α by increasing either the productivity or, if possible, the utilization of capacity.

As an example of this development we will account for some statistical observations collected from the Danish sugar industry over a period of 26 years from 1876 to 1902.

The details of the investigation must be omitted,[5] but two diagrams are submitted (Figures 1 and 2). Figure 1 shows the total amounts of

[5] Will be published by The Institute of History and Economics, Copenhagen.

money and total quantities of raw materials passing the first manufacturing stage (2 factories) of the Danish sugar industry in which raw sugar is made from beets (the second manufacturing stage is the refinery). A number of factories making white sugar directly from beets are omitted. In Figure 2 the amount of money is further expressed *per unit* of raw material passing into the stage.

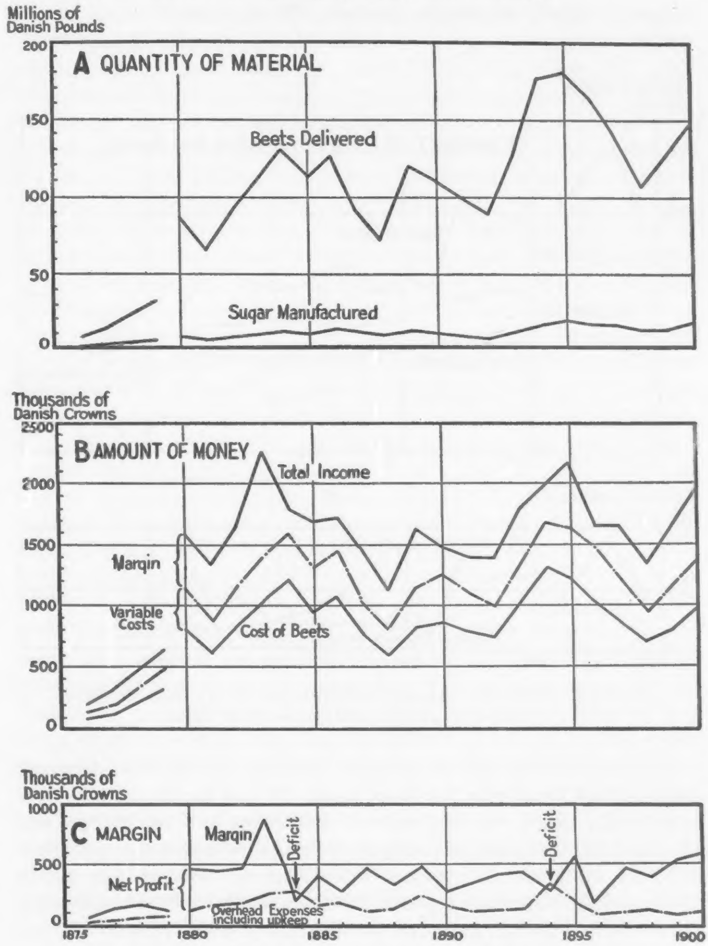


FIGURE 1.—Total production, costs, and margin in Danish raw-sugar factories: 1 factory, 1876–79; 2 factories, 1880–1900.

Of the total amount of income received (Figure 1, Section B) a certain proportion is passed on to the previous "manufacturing" stage, in this case agriculture, in payment for raw materials, i.e., the beets. This quantity is represented by the lowest curve in section B of the diagram. The difference between the top and the bottom curve is the amount of money retained at the stage in question in payment for wages, auxiliary materials, interest, upkeep, general expenses, and profit.

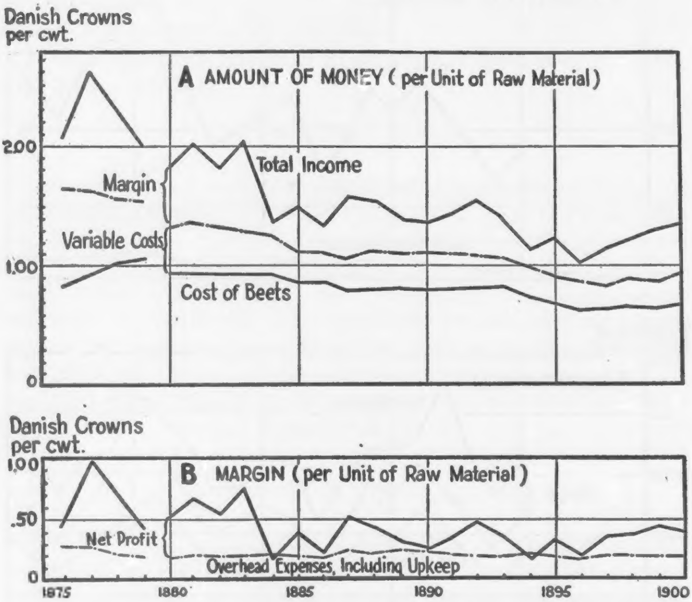


FIGURE 2.—Per-unit costs and margin in Danish raw-sugar factories: 1 factory, 1876–79; 2 factories, 1880–1900.

Between the top and the bottom curve is a third curve (dot and dash), which represents the total outlay for raw materials, *plus* what is generally called variable costs in distinction to fixed charges and overhead expenses, and covering therefore *all* outlay on raw materials, auxiliary materials, and wages. The difference between this middle curve and the upper one represents income available for fixed charges, overhead, and profits. In a separate diagram (Figure 1, Section C) this available margin has been depicted.

Figure 2 corresponds to Figure 1, but amounts of money are ex-

pressed per unit of raw material passing into the stage. In this case the whole difference between the top and bottom curves of the main diagram (Section A) is equal to the value product added per unit of raw material passing the stage or the term α of our equation.

It is quite evident that α shows an almost continuous decline over the 26 years it is possible to account for. This is a well-known fact, that is common to all industries, when technical development sets in.

Technical development, as we have seen, aims at creating a higher productivity, whereby with given and, as is generally presumed, relatively *stable* market conditions (α of our equation constant), a larger *total* income may be realized by the manufacturer, who first improves his technique. The temptation to disturb the stability of the market by lowering the unit price of the finished product is generally too big, however, for in doing so to a moderate extent the manufacturer making use of the highest technique, i.e., the highest productivity, may in the first place increase his *total* income, if he can only increase the number of units solid in a larger proportion than the unit price is lowered. But he puts a pressure on all other producers in the same line,forcing them to increase either h or q', for, as long as the desire and necessity to make a profit under the prevailing technical conditions have to be fulfilled in spite of declining prices brought about by the one producer making use of a new technique, there is no escape from the adjustment of q' or h to maintain the equilibrium, which is expressed in (7).

Once the interaction of adjustments is started it will continue as long as the productivity may be increased by any one of the three methods which previously have been outlined (Part 5). It is impossible to say that there is only one single *cause* of this adjustment, for as long as three quantities q', α, and h have to satisfy the conditions of equation (7) the change of only one will cause an adjustment of one or both of the remaining two. It may, however, no doubt be truly said, that in the modern world it was the technical improvements, which 150 years ago suddenly and vastly increased the productivity in the textile industries of Great Britain together with the technical possibility of using steam power realized shortly afterwards, which set the whole process of modern industrial development into motion, leading forward into our own times.

The mechanism of adjustment at a later stage, when decreasing α is the cause, which affects both q' and h, may be plainly seen from the diagrams of the Danish sugar industry. It will be seen from the total quantities how the available margin (Section C of Figure 1) decreases with the sharp drop in α (Section B of Figure 2) taking place in the year 1884. In 1893 considerable improvements of technique are put into use and total quantities of material increase to a high peak in

1894 and 1895. In spite of this the total quantity of income is not even as high as in 1885 and the margin available for fixed charges, profit, etc., is considerably smaller. That there is an actual deficit in 1894 is due to construction being paid out of running expenses and a new decrease in α due to the sugar crisis of 1894 (cf. Figure 2, Section B).

8. THE GENERAL ADJUSTMENT WITHIN THE SYSTEM

So far the question has been only of an adjustment, which takes place between a certain industry and the markets in which it buys and

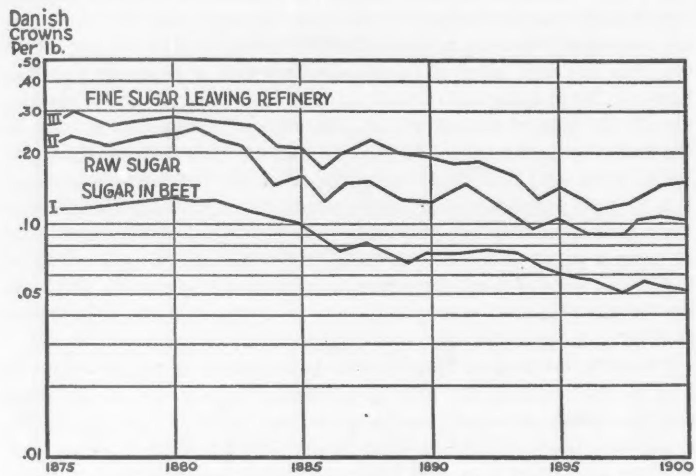


FIGURE 3.—Functionally interrelated price levels in Danish sugar industry.

sells. It is to be expected, however, that a similar adjustment is acting throughout the whole economic system, between all industries, agriculture not to be forgotton. This may be illustrated, if not for the system as such, at least for different stages of production of the same industry, by statistical observation of what might be called *functionally interrelated price levels*. The data again have been taken from the Danish sugar industry.

Figure 3 contains a representation on ratio scale showing the proportions according to which total available value product per pound of fine sugar sold to the consumers has been divided among the three manufacturing stages (agriculture, raw-sugar factory, and refinery), through which the material must pass before it reaches the consumer.

If now the linear distance between the curves of the diagram is measured in each succeeding year a table may be prepared (Table 4) and arithmetic averages determined. If then the five-year moving average is computed (columns 2 and 5) for the two sets of values, and deviations of the yearly values from the average are taken, it will be seen that these deviations show no trend over the period in question (columns 3 and 6). In other words: In spite of all technical improvements each of the three stages has in the long run had no advantage in relation to any of the others as to the amount of money retained in payment for one pound of material reaching the consumer.

TABLE 4
FUNCTIONALLY INTERRELATED PRICE LEVELS

Year	Linear Distance from Price Level I–II in mm.* (1)	5-Year Moving Mean (2)	Relation of 5-Year Mean to Average (3)	Linear Distance from Price Level I–III in mm.* (4)	5-Year Moving Mean (5)	Relation of 5-Year Mean to Average (6)
1877–78	23			32		
1878–79	19			26		
1879–80	20	103	1.145	26	137	0.988
1880–81	19	99	1.100	26	130	0.938
1881–82	22	100	1.112	27	131	0.945
1882–83	19	91	1.012	25	127	0.915
1883–84	20	87	0.968	27	126	0.908
1884–85	11	76	0.845	22	121	0.872
1885–86	15	79	0.878	25	129	0.930
1886–87	11	79	0.878	22	135	0.973
1887–88	22	85	0.946	33	148	1.067
1888–89	20	90	1.000	33	155	1.117
1889–90	20	100	1.112	35	163	1.175
1890–91	17	95	1.056	32	159	1.146
1891–92	20	95	1.056	30	152	1.097
1892–93	21	86	0.956	29	140	1.010
1893–94	17	88	0.978	26	136	0.980
1894–95	11	83	0.924	23	134	0.967
1895–96	19	82	0.912	28	132	0.952
1896–97	15	85	0.945	28	132	0.952
1897–98	20	96	1.068	27	141	1.017
1898–99	20	100	1.112	26	146	1.053
1899–00	22			32		
1900–01	23			33		
Average:		89.9			138.7	

* Measurements refer to Figure 3 drawn on a different scale from that reproduced here.

It should be mentioned, however, that in the present case the industry in question is wholly controlled by one organization and contracts for beets decided according to the market price of raw sugar on the London market.

We have already seen that in the various possibilities of technical improvements, which may call for economic adjustments, two are what may be called "self-limiting" inasmuch as it is impossible to improve the efficiency of application of either human or technical energy beyond 100 per cent. Consequently the influence of such improvements in causing adjustments of the economic structure may be expected to be relatively slight. Yet it should not be forgotten that two inventions of this nature as John Kay's flying shuttle and Hargreaves' spinning jenny, both of them improving the efficiency of the human worker only, may have contributed in an absolutely decisive manner to the development of British industry in the century following by the social forces they set free through the sudden displacement of workers on a large scale. They probably have been the impulse which started the expansion.

Compared to the social and economic instability of the modern world, however, where invention influencing productivity follows invention, always causing displacement of human workers, they were mere trifles. Further they were almost 30 years apart. The modern inventions with hardly any exception are always aiming at the increase of the technical intensity of production and as we have seen the technical conditions do not under these circumstances produce any natural limits. This is the reason why their social and economic adaptation may be expected to follow quite a different course from the social adaptation of technical improvements of efficiency. The latter will come to rest by itself. The former will continue at an ever-increasing pace.

This fundamental difference will no doubt have a profound influence on the economic conditions of a community subject to such a possibility, a consideration which so far seems to have been overlooked. It is therefore quite probable that a determination of average values of intensity for a whole economic system based on the *net* amount of energy per man-hour, which it has been possible to convert into useful work representing value product actually added, would be of considerable interest. This would be an average of the factors $i_c \cdot \eta$ of the expression (6) determined for a whole community.

L. R. Nienstaedt

Hellerup, Denmark

REPORT OF THE OXFORD MEETING, SEPTEMBER 25–29, 1936

THE SIXTH EUROPEAN MEETING of the Econometric Society was held in New College, Oxford, England, from September 25th to 29th, 1936.

The first session, on Saturday morning the 26th, was devoted to a symposium on "Mr. Keynes' System." Mr. J. E. Meade (Hertford College, Oxford) presented "A Simplified Model of Mr. Keynes' System." He supposed a closed economy; perfect competition, so that every price equals marginal cost; and only two industries, one making consumption goods and the other capital goods, having the same short-period elasticities of supply, and using no variable factor save labour and no fixed factor save equipment which does not depreciate. He further restricted his study to the short period, defined as the period in which the ratio between the output and the stock of new capital goods is so small that we can neglect changes in the stock. On these assumptions the model was given by the following conditions of equilibrium:

$$(1)\quad p_x = w\frac{dN_x}{dx}, \qquad (2)\quad p_y = w\frac{dN_y}{dy},$$

$$(3)\quad I = xp_x + yp_y, \qquad (4)\quad I = P + wN,$$

$$(5)\quad N = N_x + N_y, \qquad (6)\quad x\cdot p_x = sI,$$

$$(7)\quad i = \frac{E(P)}{px}, \qquad (8)\quad \frac{p_x\cdot K}{M - kI} = L(i),$$

where x, y, are the outputs per unit of time of capital goods and consumption goods respectively;

N_x, N_y, are the numbers of persons employed in making them; and p_x, p_y, are their prices per unit;

w is the money wage per person per unit time;

I is the total money income, and P is total money profit;

s is the proportion of income saved, so that by equation (6) savings $=$ investment;

i is the rate of interest;

$E(P)$ is the yield which is expected in each future year on a unit of equipment installed now;

K is the existing stock of equipment;

k is the proportion of money income held by the public to finance business and income transactions, so that $M-kI$ is the amount available to satisfy the precautionary and speculative motives for liquidity;

$L(i)$ is the liquidity preference function, so that by equation (8) the

ratio of the aggregate value of non-liquid assets to that of liquid is treated as a function of the rate of interest.

M, w, and s are regarded as independent variables; K and k are constant; and x, y, p_x, p_y, I, P, i, and N are dependent. Thus, given the quantity of money, the money wage-rate, and the proportion of income saved, the volume of employment is determinate. Mr. Meade proceeded to examine the conditions of stability of the system. Suppose for example that, w and s being constant, the banks vary M so as to keep i constant. If a chance increase now occurs in expenditure, and hence in employment and income, people will (a) increase the amount they try to save, and (b), since profits are raised, increase the amount they try to invest. If there is to be a return to the initial position we must be assured that as long as i is constant (a) will always outweigh (b); but that is to require that in our system an increase of employment, N, shall be possible only if i falls. The condition $di/dN<0$ is therefore taken as the criterion of stability in the present case. Differentiating equations (1) to (7), we find $di/dN<0$ if $\pi<1-l$, where $\pi=dE(P)/E(P)\cdot P/dP$ and $1-l$ is the proportion of income that goes to profits. After considering other such cases, Mr. Meade concluded by examining the effect of changes in i, M, w, and s on the short-period demand for labour, on the assumption that the system was stable.

Mr. Meade's method of testing stability gave rise to a discussion, renewed from time to time during the conference, on the possibility of making any test of stability in a purely static system. Summing up the discussion, Prof. Frisch suggested that there were two kinds of system: (1) in which an *arbitrary* monotonically increasing transformation of time is allowed; (2) in which the question: "How rapidly?" is asked, and where therefore only *linear* monotonically increasing transformations of time are allowed. In neither kind can the stability of an equilibrium be examined without bringing in dynamic considerations, but in (1) these may—to use a phrase suggested by Kalecki—be "qualitative dynamics" in the sense of Analysis Situs. A set of assumptions in (1) that is sufficient to determine an *equilibrium*, in the sense of a set of values of the variates compatible with all the conditions, will as a rule *not* be sufficient also to distinguish between a stable and an unstable equilibrium. Nearly all static theories of demand and supply, etc., are of this sort. *Supplementary assumptions* regarding the behaviour of the system are therefore necessary. In (1) the dynamic element comes in just through these supplementary assumptions that are added to those necessary to determine the equilibrium.

These supplementary assumptions may be formulated in different ways which lead to different situations so far as stability is concerned. For example, in the well-known diagram of Supply and Demand, if

one set of supplementary assumptions is adopted, the equilibrium will be stable or unstable according as $D' \gtrless S'$, and if another set of supplementary assumptions is adopted the equilibrium will be stable or unstable according as $|D'| \gtrless |S'|$, where D' and S' are the derivatives of the functions that express the quantities demanded and supplied respectively as depending on the price. It is therefore important to bring out explicitly the supplementary assumptions on which the conclusions regarding stability are based.

Contributions to the symposium by R. F. Harrod (Christ Church, Oxford), and Dr. J. R. Hicks (Gonville and Caius College, Cambridge), have been published in ECONOMETRICA.[1]

On Saturday afternoon, Professor Frisch spoke on "Macrodynamic Systems Leading to Permanent Unemployment." He pointed out that the possibility of permanent unemployment even where there was no restriction of the supply of labour, to which Mr. Keynes had called attention in his recent work, had during the last ten years been brought out in macrodynamic studies by Amoroso, Vinci, Roos, Frisch, Kalecki, and Tinbergen. To suggest the nature of this kind of analysis, he gave a simple example drawn from lectures which he had delivered at Oslo in the winter of 1933–34. Assume a community in which only one commodity, say wheat, is produced, the only capital is a stock of wheat, and there are only two kinds of person—labourers who consume all they get, and entrepreneurs who save all they get. Put

x = the output of wheat per unit of time;
a = the labour coefficient in hours of work per bushel of wheat;
w = the wage-rate, in bushels per hour;
K = the stock of wheat (capital), in bushels;
π = the gross profit $(=1-aw)$, in bushels per bushel;
κ = the depreciation rate, in % per unit time;
$\dot{K} = dK/dt$ = the growth rate of capital, which is the same thing as total net profit per unit of time.
$\lambda = \dot{x}/\dot{K}$ = sensitiveness of expansion.

If we assume w, a, κ, and λ constant (it is actually necessary to assume only that they move more slowly than the other terms), we are left with two equations in two variables x and K:

$$\dot{K} = \pi x - \kappa K$$

$$\lambda = \frac{\dot{x}}{\dot{K}}.$$

[1] R. F. Harrod, "Mr. Keynes and Traditional Theory," ECONOMETRICA, Vol. 5, Jan., 1937, pp. 74–86; J. R. Hicks, "Mr. Keynes and the Classics; a Suggested Interpretation," *ibid.*, Vol. 5, Apr., 1937, pp. 147–159.

The following time-shapes ensue:

$$x_t = \bar{x} + ce^{\gamma t}$$
$$K_t = \bar{K} + Ce^{\gamma t} \text{ } (c, C, \overline{X}, \text{and } \overline{K} \text{ constants}),$$

where $\gamma = \lambda\pi - \kappa$. Now if γ is negative, x_t and K_t approach stationary values, and since a is assumed constant, this means a stationary amount of employment, which need not be full employment. After showing how the system might be modified and developed in a more realistic direction, Professor Frisch remarked that today, when there are already some 10 or 15 fully-developed mathematical systems of some plausibility, the task was not so much to develop new systems as to test different systems against the facts.

At the same session Dr. J. Tinbergen (Rotterdam and Geneva) introduced a discussion of "Dynamic Equations Underlying Modern Trade Cycle Theories." He offered a survey of the different attempts that have been made to describe the endogenous movements of an economic system by a complete system of equations, i.e., by a system showing as many equations as variables. To fix ideas he began with a description of his own attempt to describe American cycles after the war by means of a much simplified system of some 12 equations. The simplification is made possible notably by the facts that the American balance of trade is very small in comparison with national income, and that the influence of interest rates on the volume of investment activity seems to have been very small in the period considered. By choosing suitable units and zero points for his variables, which were measured as deviations from a moving equilibrium. Dr. Tinbergen succeeded in obtaining a system of linear homogeneous equations. The aim of these equations was to indicate the "direct causal relationships" existing between the variables as long as no exterior forces are acting. One of the equations, for example, is $b_t = \beta Z_{t-1}$. Here b_t is the deviation from its equilibrium level of investment activity in period t (the length of the unit period is four months); Z_{t-1} is the deviation from its equilibrium level of total nonlabour income in period $t-1$; and β is a constant. Another of the equations is $E_t = \epsilon Z_t$, where E_t is non-labour income distributed (as distinct from nonlabour income accruing, Z_t) as a deviation from its equilibrium value, and ϵ is another constant. The constants used are "reaction coefficients," whose values depend largely on institutional, technical, and natural factors, and as a rule change only slowly. As a solution of the system of equations shows, they determine the nature of the endogenous movements through which the system can go. In many cases their values can be determined by multiple or simple correlation analysis, so far as exterior forces can be considered as accidental. The question can then be put, what

changes in these coefficients are necessary in order to obtain a system that will not show heavy fluctuations: this is the problem of business-cycle policy.

Comparing his system with others, Dr. Tinbergen restricted himself to the mathematical representations of Mr. Keynes' and Prof. von Hayek's theories given by Mr. Meade and Mr. Thompson. He saw much resemblance between Mr. Keynes' system and his own, but Mr. Meade's equations are static, not dynamic, and in consequence speculative gains do not enter into them; but in Dr. Tinbergen's opinion such gains have played an important part in the last American cycle. Mr. Thompson's representation of Prof. von Hayek's theory was dynamic, but in a rather oversimplified way, only one lag being introduced; and a number of the hypotheses underlying this system seemed to Dr. Tinbergen rather arbitrary.

In replying to discussion, Dr. Tinbergen put forward a number of considerations to justify the taking of a single system to fit the course of a sequence of cycles: (1) It was insufficiently realised that the mechanism which yields a cumulative process may also yield an explanation of the whole cycle: there is then no need to introduce special causes of turning points, for these are implicit in the single mechanism. (2) Though it is true that the system is disturbed by shocks, the shock is none the less transmitted *through* the system. When the commercial records show a shock of temporally well-defined impact, as in 1926 in England, the figures of that time can be omitted from the data used to determine empirically the relations postulated by theory. (3) It was suggested that a fuller account might be obtained if different systems were fitted to different phases or cycles; but it was difficult to get sufficient data for the short period, and when he had tested the fit in different periods, he had not found any systematic difference according to the phase of the cycle. In general, it is better to include more variables than to resort to more complicated functions.

In a subsequent colloquium, Prof. Frisch developed an ideal programme for macrodynamic studies: A. *Theoretical inquiry.* (1) Define your variables. (2) State the structural relations which you suppose to exist between the variables. (3) Derive a number of confluent relations, which lead to confluent elasticities, showing the response of one variable in a certain sub-group to another when all the rest are held constant. (3a) Use these relations for reasoning about variations compatible with the subsystem. (3b) Consider the response of the system to exogenous shocks: a dynamic analysis leading to criteria of stability. (3c) Consider how the whole system will evolve in time. B. *Statistical inquiry.* (4) Obtain some final equations. A final equation is a confluent relation which is reduced to its smallest degree of freedom,

and in which the coefficients have a statistically uniquely determined meaning. Never try to fit to the data anything but a final equation. (5) One also may scrutinise the data, and derive empirical formulae by the statistical technique now known as "confluence analysis." In particular a final equation may be tested in this way. (6) If the final equation contains only one variable, and is linear, construct the corresponding characteristic equation and consider its roots. They will determine the time-shape of the evolution that would ensue if the system were left to itself. When we have found in this way how the system *would* proceed through time, we do not expect actual history to move like that, for this history is affected by a stream of erratic shocks. The actual time-shape will now be a weighted average extending over the shocks, the weights of the average being those given by the system as it would proceed in isolation. (7) Fundamental inversion problems: (a) to determine the system of weights from a given time-shape; (b) to determine the shocks. (8) Attempt a forecast using the weights determined by the inversion, and assuming—in the absence of better information—the future shocks to be zero.

On Sunday morning, the 27th, Mr. R. G. D. Allen (London School of Economics) spoke on "The Assumptions of Linear Regression." Mr. Allen first pointed out that the linear regressions obtained by the method of least squares are not appropriate to the problem of deducing the "best" linear relation between observed variables; and he proposed the following procedure. Let us assume that the true relation is linear, but that in the observations this relation is obscured by displacements in both x and y. Further let us assume (1) that the sums of the displacements are zero, (2) that there is no correlation between the displacements of x and y and their true values. If ξ, η, are true values (measured from their means), then

$$\eta = m\xi,$$

$$x = \xi + u, \tag{i}$$

$$y = \eta + v. \tag{ii}$$

By (1), we can take u and v as measured from their means. By (2), we have Mean (ξu) = Mean $(\eta v) = 0$ = Mean (ηu) = Mean (ξv). In Frisch's notation,

$$\lambda_u = \frac{\sigma_u}{\sigma_x}, \qquad \lambda_v = \frac{\sigma_v}{\sigma_y}.$$

Then squaring (i) and (ii) and eliminating ξ, η, we get

$$m = \pm \frac{\sigma_y}{\sigma_x}\sqrt{\frac{1-\lambda_v^2}{1-\lambda_u^2}}. \tag{I}$$

Similarly, multiplying (i) and (ii),

$$
\text{(II)}\qquad
\begin{aligned}
r_{xy} &= m\frac{\sigma_x}{\sigma_y}(1-\lambda_u^2) + r_{uv}\lambda_u\lambda_v \\
&= \pm\sqrt{(1-\lambda_u^2)(1-\lambda_v^2)} + r_{uv}\lambda_u\lambda_v.
\end{aligned}
$$

These results are independent of scale. As particular cases, we have from (I) and (II),

$$
\text{(1)}\qquad \lambda_u = 0, \qquad m = r_{xy}\frac{\sigma_y}{\sigma_x} = m_1,
$$

$$
\text{(2)}\qquad \lambda_v = 0, \qquad m = \frac{1}{r_{xy}}\cdot\frac{\sigma_y}{\sigma_x} = m_2,
$$

$$
\text{(3)}\qquad r_{uv} = 0, \qquad m = \frac{r_{xy}}{1-\lambda_u^2}\,\frac{\sigma_y}{\sigma_x} = \frac{1-\lambda_v^2}{r_{xy}}\cdot\frac{\sigma_y}{\sigma_x},
$$

$$
\text{(4)}\qquad r_{uv} = \pm 1.
$$

On putting

$$
\frac{v}{u} = \tan\beta,
$$

$$
m = \frac{\sigma_y\cos\beta - r_{xy}\sigma_x\sin\beta}{r_{xy}\sigma_y\cos\beta - \sigma_x\sin\beta}\cdot\frac{\sigma_y}{\sigma_x}.
$$

Take $r_{xy}>0$. Then m lies within the range (m_1, m_2) in case (3) and also in case (4) if $\tan\beta<0$. But m lies outside (m_1, m_2) in (4) if $\tan\beta>0$. The elementary regression values, m_1 and m_2, are, in fact, not necessarily limits for m if there is correlation between the displacements.

At the same session, Dr. Jerzy Neyman (University College, London), gave a "Survey of Recent Work on Correlation and Covariation." He considered three classes of problem: I. The classical mathematical problems of stochastic dependence, correlation, regression, testing statistical hypotheses, and estimation. Here the problem is to deduce the consequences of certain hypotheses which are not subject to question whenever they are not contradictory; and to test the hypotheses by comparing their consequences with the available observations. II. Applicational problems, in which the task is to adjust a system of hypotheses, mathematically expressed, in such a way that their consequences will conform with the available observations. Here there are two paths of approach, the empirical, and the a priori The difference between them "may be emphasised by the comparison with two different phases in the history of astronomy. The first phase is

that of Ptolemy to Kepler. Ptolemy laid down the principle that mathematicians should endeavour to represent all the celestial phenomena by uniform and circular motions. This principle was followed roughly till Copernicus, and was removed by Kepler. But all the effort in this phase consisted in *guessing* the appropriate formula and in adjusting the numerical coefficients so that it might fit the observations. That is what is being done in the empirical approach to social and economic phenomena, e.g., in the analysis of time series, which we try to split into trend, business cycles, seasonal variations, etc. In astronomy the new era began with Sir Isaac Newton, and with his set of hypotheses concerning not the functions representing the observable facts, but the machinery which may have produced those facts. This is what may be termed an a priori construction." III. Mathematical problems marking a new era: the theory of random differential and integral equations, and the corresponding problems of testing hypotheses and estimation. In I and II we *assume* the hypotheses, and apply them; but in III we *test* them. Here Dr. Neyman referred again to the analogy of astronomy. "The success of the work of Newton and others is due largely to the existence of the special branch of mathematics (calculus) suitable for dealing with the problems treated. Similarly, the success of social and economic dynamics depends on the existence of the appropriate branch of mathematics. This branch does not exist in a ready form at present." The phenomena we observe are affected by random fluctuations: if then "we want the formulae describing these phenomena, and want to deduce them from some hypothesis concerning their machinery, the ordinary calculus could not be sufficient—we must deal with differential and integral equations which are in a sense random or stochastic. As far as I know, the honour of noticing the necessity of such 'stochastic calculus' must be ascribed to Harold Hotelling, who was the first to consider stochastic differential equations, in his paper on 'Differential Equations Subject to Error, and Population Estimates,' *J. Am. Stat. Ass.* 1927. A systematic study of the problem was begun by Serge Bernstein, who read a paper on the subject before the International Mathematical Congress in Zürich in 1932, and has since published a series of important papers in the *Reports of the Leningrad Academy* and in Vol. V of the *Annals of Stekloff Phis. Math. Institute.* I have the impression that these works mark the approaching new era in the study of problems which are now classified as those of correlation and covariation.".

In the course of his remarks under III, Dr. Neyman developed an account of what may be called "the Neyman-Pearson theory of testing hypotheses," of which the following summary has been prepared by Professor Frisch:

Let Ω be a set of admissible hypotheses. The content of Ω is made up by selecting certain hypotheses out of a still larger conceivable range of hypotheses. Once a hypothesis is admitted into the admissible group Ω, it is considered on a par with all the other hypotheses in Ω. In other words, there does not exist any "a priori probability" by which it is possible to discriminate between the hypotheses contained in Ω.

An observation E is given. E consists of a measurement of each of n variates $x_1, \cdots, x_n$. The space $(x_1, \cdots, x_n)$, is called the sample space. The problem is to draw conclusions as to which one of the hypotheses in Ω is true, when the only information available consists of E (and the construction of Ω).

The Neyman-Pearson approach is essentially a "method of rejection." That is to say, for any given hypothesis h_0 of Ω, one tries to answer the question: Shall this hypothesis be rejected, or shall it not be rejected? The latter alternative may be interpreted as "the hypothesis may be true." By successively applying this procedure to several hypotheses a certain number of them may be rejected and so the range be limited. This process may be looked upon as one of using E to restrict the content of Ω from what it was originally.

The starting point of the method is to associate with each hypothesis h_0 a certain region ω_0 of the sample space and to make the convention that the hypothesis shall be rejected when and only when E falls in ω_0. The latter event will for shortness be denoted $E \subset \omega_0$. The construction of ω_0 is thus a definition of a method of rejecting h_0, using the information E. The region ω_0 is called the critical region for h_0. The essential point in the theory consists in providing definite rules for the construction of a region ω_0 associated with each h_0.

This construction is based on two kinds of probability, namely the probability of rejecting h_0, if h_0 is true, and the probability of accepting h_0, if some other hypothesis h is true. The first is called the probability of errors of the first kind P_I, the latter is called the probability of errors of the second kind P_{II}.

It is assumed that to each hypothesis h there corresponds a certain probability law of E, that is to say, a theoretical distribution of the possible observations $(x_1, \cdots, x_n)$ in sample space. If such a probability law is given for each hypothesis h and if ω is any region in sample space, $P\{E \subset \omega/h\}$ "the probability that E shall fall in ω if h is true" is also given. In terms of this probability law the probabilities P_I and P_{II} may be expressed as follows:—

(1) $$P_I = P\{E \subset \omega_0/h_0\},$$

(2) $$P_{II} = 1 - P\{E \subset \omega_0/h\}.$$

It will be seen that P_I depends only on h_0, while P_{II} depends both on h_0 (because the critical region ω_0 associated with h_0 occurs in formula (2)) and on h.

The method of constructing the critical region can in short be said to consist of "making both probabilities P_I and P_{II} as small as possible." This statement needs of course to be made more precise.

"To make P_I small" is postulated to mean that P_I is to be made equal to some conventionally chosen number, for instance $\alpha=0.1$. Some such conventional limit must of course be fixed, because, by reducing the content of ω_0, P_I can always be made as small as we please.

This first postulate may be satisfied in an infinity of ways. To achieve the determination of ω_0 the further condition is added that P_{II} is to be made a minimum with given h. When ω_0 is selected in this way, the test obtained is said to be "the most powerful test of h_0 with respect to h." Similarly the region ω_0, which is defined in this way, is called "the best critical region of h_0 with respect to h."

Sometimes it happens that the region ω_0 defined by the above principle is independent of h (while depending of course on h_0). In this case it is said that a uniformly most powerful test of h_0 exists. It is, however, only in exceptional cases that such uniformly most powerful tests exist. For instance, if the various hypotheses h can be characterised by different values of a continuous parameter θ (or by a number of such parameters $\theta_1, \theta_2, \cdots$) and if certain derivatives exist, then uniformly most powerful tests do *not* exist.

Since uniformly most powerful tests do not as a rule exist, some other criterion must be found to replace them.

Consider the probability

$$(3) \qquad \beta(h, h_0) = 1 - P_{II} = P\{E \subset \omega_0/h\}.$$

This is the probability of rejecting h_0 if h is true. This probability will of course vary with h. Let us consider this variation with h if h_0 is kept fixed. The desideratum is of course to have a function β which is the smallest possible for that particular value of h which is equal to h_0. Indeed, this is the probability of rejecting h_0 if h_0 is true, in other words, it is the conventionally chosen figure α. But outside the particular point $h=h_0$ we want the value of β to be as high as possible. The ideal would be a discontinuous function β which is equal to the small number α for $h=h_0$ and to unity for all other values of h. In general this cannot be arranged, but if the various hypotheses h may be characterised by a single parameter θ, one may try to arrange it so that β, considered as a function of θ (for a given h_0), has its minimum minimorum for that particular value $\theta=\theta_0$, which corresponds to the hypothesis h_0, and further such that β increases as rapidly as

possible as θ moves towards the right or towards the left from the point θ_0.

If the test is formulated in such a way that β does not have a minimum minimorum in the point θ_0, but reaches still lower values in some other points, the test is called *biased.* In this case h_0 will be rejected more frequently (in a number of cases in each of which an observation E is given) when it is true than when certain other hypotheses (namely those for which β is lower) are true. If the test is such that β has a minimum minimorum for $h=h_0$, the test is called *unbiased.*

At the same session a paper by Dr. N. Georgescu (Bucarest), on "Scatter Analysis," was read by title.

On Sunday afternoon, 27th, Dr. A. P. Lerner (London School of Economics) gave a survey of recent work on International Trade and Transfer, saying in part: "Recent writers on International Trade have been troubled by the difficulty of fitting the theory of the subject into the general framework of economic doctrine, and this is reflected in the whole tenor of their work. The difficulty arises from the peculiar historical development of International Trade Theory. An initial precocity at the beginning of the nineteenth century, when the Law of Comparative Costs *partly* freed it from the labour theory of value, put it ahead of the other branches of economics, but, not unnaturally, was followed by a retardation of further development while the rest of economic theory achieved a more complete emancipation. To this situation some writers, like Ohlin and Iversen, react by rejecting the whole of the classical structure and building anew an Interdependence Theory of international trade on Walrasian lines. Others, like Haberler and Viner and (more intensively) Harrod, defend the classical real-cost analysis by a process of benevolent interpretation. Ohlin appears to be too much enamoured of the Walrasian scheme to be willing to admit that real (if incomplete) results can be obtained only by partial analysis. Instead of passing from a Walrasian preface to a body of particular analyses, he tries to deal with everything. The result is an exercise in one-thing-at-a-time analysis attempting to be everything-together analysis by jumping very quickly from one thing to another. Ohlin makes some concessions to the need for partial analysis by simplifying the whole scheme somewhat. In doing this he falls between two stools, achieving neither complete generality nor realistic particular results, and arrives at some false conclusions in the theory of the equalisation of factor prices by international commodity movements. The main merit of his method lies in its suggestiveness of interconnections that might be overlooked.

"In the theory of Transfer, the use of the concept of *the transfer of buying power* developed by Ohlin (following Bastable and in a sense

Ricardo) and the deprecation of the discussion of elasticities of demand, is dangerous procedure. It is possible by a fairly simple diagrammatic scheme to show that the terms of trade will move in favour of or against the paying country according as the elasticities of supply of exports and imports are greater or less than the elasticities of demand, all measured in terms of domestic goods. Because of the possibility of the movement of factors of production between domestic and export goods but not between domestic and import goods, there is a presumption that the elasticities of supply are greater, so that the terms of trade are *more likely* to move against the paying country. Discussion on this topic has been confused by the entry of irrelevant problems and prejudices.

"Harrod keeps to a real-cost analysis in order to be able to discuss the gain from international trade. It can be shown that all his results are obtainable without the objectionable and difficult real-cost analysis, and can be demonstrated much more simply by the use of the indifference-curve technique.

"Another important case of a traditional bias is the belief that taxing the foreigner by import or export duties is a practical impossibility. Haberler makes great use of this. But the application of the theory of monopoly gain shows that there is an ideal tax on every import or export commodity of not less than the inverse of the elasticity of supply or demand from the point of view of the country. (More generally,

$$\frac{e_p/e_c - 1}{e_p - 1},$$

where e_c is the elasticity from the point of view of the country and e_p is the elasticity from the private point of view of importer or exporter. In perfect competition this reduces to $1/e_c$.) This becomes negligible only if *all* elasticities are of a very high order, an assumption for which there appears to be no basis."

At the same session, Dr. J. Wiśniewski (Warsaw) spoke on the "Measurement of Real Income," saying in part: "There can be two approaches to the index numbers of prices, viz., (A) the functional approach, and (B) the behaviouristic (by Frisch termed atomistic) approach. (A) I. Homogeneous communities. Their members differ from each other (a) by incomes, (b) by individual tastes. The prices for which they buy goods are the same for all individuals in a given situation. Within a homogeneous community average Engel functions are considered as independent of income. Suppose functional index numbers are known. A certain individual enjoyed in situation 0 an income x_0, and in situation 1 an income x_1. By means of functional

indexes we can state that an income x_0 in situation 0 is equivalent to an income x_0' in situation 1, and that an income x_1' in situation 0 is equivalent to x_1 in situation 1. We should thus be involved in a paradoxical situation: the ratio of real incomes enjoyed by our individual in situations 1 and 0 would be either $x_1':x_0$ or $x_1:x_0'$. A way out of this difficulty would be to measure price changes along an 'historical path' according to the money income the individual actually enjoyed in each moment. II. Nonhomogeneous communities. New difficulties are present here. If prices are not the same for individuals at different income levels, the proportionality of money and real incomes cannot be assumed. Moreover, if average Engel functions are not the same for the subgroups of the community the functional approach is likely to fail. (B) The behaviouristic approach may best be represented by Divisia's method (*Rev. d'Econ. Pol.*, 1925, p. 999 seq.). Conditions may be specified for the 'true' index computed by Divisia's method, to be contained within the familiar formulae of Laspeyres and Paasche."

At the same session Mr. Colin Clark (Cambridge) made a survey of "Measurements and Comparisons of National Income." He offered the following classification of available estimates of national income:

(1) Prepared by the government itself from official sources: Germany, U.S.A., U.S.S.R.

(2) Semi-official, compiled by persons who are members of government departments: Australia, Canada, France.

(3) Made by academic institutes using mainly official material: Denmark, Holland, Hungary, New Zealand, Sweden.

(4) Compiled by banks: Greece, Switzerland.

(5) Privately compiled from official sources: Great Britain, Irish Free State, Japan, South Africa.

(6) Infrequent: Austria (Konjunkturinstitut, 1928), Belgium (Baudhuin), India, Italy, Poland.

(7) None: Czecho-Slovakia, S. America, Spain.

Mr. Clark then displayed the results of some recent inquires into the general level of real income, and the relative importance of industry, agriculture, and other activities, as sources of income, in different countries.

On the morning of Monday the 28th, Mr. Trygve Haavelmo (Oslo) spoke on "Confluent Relations as a Means of Connecting a Macrodynamic Subsystem with the Total System," saying in part: "A partial macrodynamic system is a study of a certain *group* of economic factors in business cycles. We shall here be concerned with problems arising when we try to make such systems determinate. We begin by setting up certain *structural* equations. These are equations which say something about *how* and *why* certain factors influence certain

other factors: they express a theory. Suppose that we have n variates, and can find only $n-m$ structural equations. Three ways may be suggested for getting the m additional equations: (a) To 'press more theory into the subsystem,' viz., to add more theoretical structural equations. This is not satisfactory, because it would in all probability be too abstract and unrealistic. (b) To suppose known the analytic expression for the time-shape of one or more variates. This is begging the whole question: it is curve fitting, not an attempt at explanation. Furthermore, a plausible analytic expression for such a time-shape is hard to find. (c) To use some empirically determined confluent relations. This means that we look at the observable result, namely the time-shape of the variates studied, and try to describe their time-shape confluency. By this means we get a *realistic* connection with the total system. It is realistic because the shape of the observed confluency is derived from the facts, and, if our theory is at all reasonable, this time-shape must be characteristic for the machinery; it must be the result of forces and actions in the total system which we are studying." As an example, Mr. Haavelmo presented a system dealing with stock- and bond-market variates, in which the confluent relations used were simplified expressions for the facts underlying the Harvard *A-B-C* lags. A numerical study of this may appear in ECONOMETRICA.

In the discussion on Mr. Haavelmo's paper Dr. Marschak raised the question of the distinction made between the structural and the confluent relations, suggesting that they differed only with respect to the *source* of the data, both being ultimately empirically determined. Professor Frisch replied that (a) there was a difference in principle between the two kinds of relation, in that any coefficient in a structural relation might be changed *institutionally* without necessarily entailing a change in the other structural relations. A similar property did not belong to the confluent relations, whose coefficients are just consequences of the coefficients in the various structural equations, and therefore change with these. Another difference was (b) that in the case of confluent relations (for instance, Harvard lags), one was asking: What *is* in fact the situation? While in the case of the structural relations one was asking: What *would be* the situation with regard to one variate *if* certain other variates, rates of change, etc., had such and such values?

At the same session Dr. A. Bijl (Amsterdam) spoke on "Wage Subsidies and Unemployment." He first contended that though a general wage reduction would increase employment in the first instance, its ultimate effect is uncertain, as we do not know whether the increment of profits to which it gives rise on impact will be disbursed or hoarded. A wage subsidy, however, the funds for which were sup-

plied by a tax proportional to profits, would have the advantage of a wage reduction without the uncertainty: for since it could be adjusted so as to leave total profit unaltered, there would be no possibility of increased hoarding, while like the general wage reduction it would lower marginal costs. Dr. Bijl went on to consider Pigou's reasons for distrusting wage subsidies, and to argue that if his scheme was so arranged that total tax exceeded total subsidy in booms and fell short of it in depressions, there would be a dampening effect on the cycle.

The third paper of the session was given by Dr. Johan Åkerman (Lund), who spoke on "Premises of Trade-Cycle Theory." He suggested that economic theory had moved through three stages: (1) in which it was concerned only with equilibrium, crises and changes being left out of account; (2) in which to a unified equilibrium theory was added a heterogeneous trade-cycle theory; (3) in which a unified theory of time changes is to be developed. Present theories of the cycle might be divided into four groups: (i) Theories of equilibrium (leading to a stabilization policy): Fisher, Cassel, Keynes 1930, Hayek 1931. (ii) Theories of anticipation (leading to policy of mitigation of monetary risk): Knight, Myrdal, Lindahl, Keynes 1936. (iii) Theories of cumulative change (leading to a cycle-smoothing policy): Wicksell, Schumpeter, D. H. Robertson. (iv) Hypothesis of activity periods (leading to policy of regulating interdependence of activity periods): cf. J. Åkerman, *Om det ekonomiska livets rytmik*, Stockholm, 1928, and *Ekonomisk Kausalitet*, Lund, 1936. In stage (3) of the development of theory, all monocausal theories of the cycle must be abandoned, and the task is to develop an account of the holistic movement containing the interplay of all its elements. As examples of such elements, Dr. Åkerman considered the different types of activity period, through which economic changes take their course. There is, for example, the commercial period, of say one year, in which the consumer and trader make their dispositions according to the theories of choice and of marginal profitability; the industrial period, of from one to ten years, in which enterprise sets itself to the task of capital construction; and the investment period of more than ten years, in which a developing country or region applies itself to long-period construction. Trade-cycle analysis must further take into account the institutional setting; the prevailing secular tendencies and seasonal variations, distinguishing the active from the passive; and the particular type of cycle and of faulty investment in the different activity periods. Because of the cumulative processes in the rising and declining phases of the cycle, the interference between activity periods of different span during the rising phase, and the impossibility of determining a capacity point, we can usefully analyse only the trough: "trade-cycle analysis proper

is thus essentially an analysis of the conversion of the depression into revival."

On the afternoon of Monday, the 28th, Dr. Hans Staehle (I.L.O., Geneva) presented "A Method of Comparing the Cost of Living in Different Situations, with its Application to Poland (1927), Estonia (1925), and Finland (1921)." He first described his measure, D, of dissimilarity in consumption. Let $q_1', q_1'', \cdots$ and $q_0', q_0'', \cdots$ be the quantities consumed of the articles entering into the expenditure budget of any consumption unit (a "man," a family, or a nation), in two situations 1 and 0; and let $p_1', p_1'', \cdots$ etc., be the corresponding prices. For each commodity consider the absolute difference between q_1/q_0 and the weighted average $\sum q_1p_0/\sum q_0p_0$, and express this difference as a proportion of the average, viz., write

$$\frac{q_1}{q_0}\cdot\frac{\sum q_0p_0}{\sum q_1p_0} - 1.$$

Taking these relatives without regard to sign, weight each by the proportion of total outlay spent in situation 0 on the commodity concerned; and form the weighted average D,

$$D = \sum \left| \frac{q_1}{q_0}\cdot\frac{\sum q_0p_0}{\sum q_1p_0} - 1 \right| \frac{q_0p_0}{\sum q_0p_0} = \sum \left| \frac{q_1p_0}{\sum q_1p_0} - \frac{q_0p_0}{\sum q_0p_0} \right|.$$

Now it is an empirical fact that when D is computed for a number of food articles bought by families of different income per "man" in one and the same country, it increases with the difference in income. When the comparison is made between different countries, or periods, one basic set of q_0's, say $\bar{q}_0$, being compared with the q_1's of families having successively higher money incomes per "man," then generally D first diminishes, passes through a minimum which is >0, and then increases; and if now the basic set is taken from a family of income higher or lower than that belonging to the $\bar{q}_0$ first taken, the new minimum in D is found for a family in 1 of income higher or lower than that which gave the minimum for $\bar{q}_0$. Now assume (1) that all families in the same situation have the same tastes, prices, etc., so that with equal incomes they would show identical budgets; (2) that each D is due to two factors, whose effects are additive—(a) differences between the two *situations* in tastes, prices, etc., (b) differences in the psychic *incomes* per "man" of the families compared. Then provided (α) that the minimum values of D are equal for all choices of $\bar{q}_0$, and (β) that if when we compare all the q_1's and $\bar{q}_0$ a minimum appears at $\bar{q}_1$, then when we compare all the q_0's with $\bar{q}_1$ a minimum appears at $\bar{q}_0$: we may conclude that (A) the minimum value of D expresses the *irreduc-*

ible dissimilarity due to differences in tastes, prices, etc.; (B) between two money incomes for which D is a minimum, the dissimilarity due to differences in psychic income is zero. The ratio of those two money incomes may then be taken as an index of the cost of living for a certain level of psychic income.

These relations may be represented in a three-dimensional diagram. On the horizontal axes plot money incomes per "man" in situations 0 and 1, on the vertical axis plot the values of D computed from the q_0 and q_1 sets belonging to the money incomes. The data then satisfy conditions (α) and (β) above if the surface shows a level-bottomed valley running obliquely through the income-income field. Dr. Staehle displayed such diagrams constructed from the budgetary materials obtained in Poland, Estonia, and Finland.

In discussion, Dr. Wiśniewski expressed a doubt concerning the applicability of the D-test to comparisons of different countries or distant years. The minimum value of D is given for the most similar structure of expenditure, and such similarity is quite consistent with differences of real income. For example, if we have two groups facing the same price situation, it will probably be allowed that real (not psychic) income is proportional to money income; but if the one group were composed of clerks and the other of labourers, the proportion of income spent on food being higher among the latter, D would not necessarily show a minimum for equal money incomes. In reply, Dr. Staehle stated that he had worked on Swedish budgets of different class groups, and found by the D-test that equal *money incomes* were equivalent, and not the incomes out of which equal food expenditures were made. Mr. Allen suggested that Dr. Staehle's condition (α) was already implicit in his assumption (1).

At the same session Prof. Ragnar Frisch (Oslo) spoke on "Determinateness and Indeterminateness in the Measurement of Money Flexibility."

Using the same terminology as in his January, 1936, ECONOMETRICA Survey, he considered three questions: (1) Under what conditions on the indicator will there be expenditure proportionality? (2) When will the money flexibility θ be an indifference function? (3) Can θ be used to define equivalent expenditures even if it is not an indifference function?

Let $W(q^1, \cdots, q^N)$ be an indicator, $E_t = p_t^1 q_t^1 + \cdots + p_t^N q_t^N$ the expenditure in the t-situation, i.e., along the t-expansion path, p being prices, q quantities. Let $W^h = \partial W / \partial q^h$. Consider (1) $w_t = dW/dE_t$ (in the following d always indicates a differential along the path). Then (2) $W^h = w_t p_t^h$ (w_t independent of h). The *money flexibility* is (3) $\theta_t = d\log w_t / d\log E_t$. It is invariant for a linear transformation of W.

The *passus coefficient* $\mathfrak{E}=\mathfrak{E}(q^1, \cdots, q^N)$ defined by (4) $q^1W^1+\cdots+q^NW^N=\mathfrak{E}W$, *characterises the nature* of W. For instance, W is homogeneous of degree $\mathfrak{E}$ if, and only if, $\mathfrak{E}=$constant. Along t write $\mathfrak{E}_t$. Then by inserting (2) into (4) we get the *fundamental identity* (5) $w_tE_t/\mathfrak{E}_t=W$ (independent of t). Hence (6) $\mathfrak{E}_t=d\log W/d\log E_t$ (the *logarithmically* marginal money utility) and (7) $d\log E_t/d\log E_0=\mathfrak{E}_0/\mathfrak{E}_t$, 0 and t being two paths. Logarithmic differentiation of (5) gives

$$(8) \qquad 1+\theta_t = d\log J_t/d\log E_t = \frac{dJ_t}{dW}, \qquad J_t = \mathfrak{E}_tW.$$

Since expenditure proportionality is characterised by (9) $d\log E_t/d\log E_0\equiv 1$, (7) gives immediately the proposition: I. *Necessary and sufficient for expenditure proportionality is that the passus coefficient* $\mathfrak{E}$ (*or, which amounts to the same,* J) *is an indifference function.* Further, the amount of disproportionality can be measured by (7).

From (8) we get: II. *If the passus coefficient* $\mathfrak{E}$ *is an indifference function, the flexibility* θ *is also an indifference function.* Hence: III. *If there is expenditure proportionality, the flexibility* θ *must be an indifference function.*

Now consider the case of *independence,* i.e., W is of the form (10) $W=U+V$, U depending on some of the q's and V on the others. In this case the left-hand member of (4) is of the form $H+K$, H depending only on the q's from U and K only on those from V. If we also have expenditure proportionality the right-hand member of (4) is an indifference function, i.e., of the form $J(W)$. Differentiating first with respect to a q in U and then with respect to one in V, we see that the second derivative of J is zero. Hence (11) $J=A+BW$, A and B constants, so that by (8) $\theta=B-1=$constant. In other words: IV. *If there is not only expenditure proportionality but also independence, the flexibility* θ *must be not only an indifference function but even a constant.* This may be called Burk's proposition because it was first proved by Abram Burk of Harvard (by another and much more elaborate method). If $\mathfrak{E}$ is an indifference function, the passus coefficients of all the W^h are the indifference function $\mathfrak{E}-1+d\mathfrak{E}/d\log W$; hence all the partial rates of substitution W^h/W^k have a passus coefficient of zero, i.e., they are homogeneous functions of degree zero.

II only gives a sufficient condition for flexibility indifference. More generally θ is an indifference function if the right member of (4), namely $\mathfrak{E}W=J$, is of the form (12) $J=F(W)+G$ where $F(W)$ is an indifference function and G a function that is constant along any path, for instance a homogeneous function of degree zero in $W^1, \cdots, W^N$. If G is an indifference function, we get back to the case covered by II.

An important question is whether there exist functions of the form (12) where G is not an indifference function and where the corresponding W is of the independence form (10). So far he had not been able to construct such functions, but was inclined to believe they exist. If so, and if the indicators we are likely to meet in reality can be assumed to have approximately this property, a method of making price comparisons between *structurally different markets*, i.e., markets with different commodities, for instance distant countries, would be available. Indeed, the flexibility would be statistically observable in each country (by the methods he has previously developed) and equal values of this parameter could plausibly be taken to define equivalence of expenditure, because if this method were applied to two situations in the *same* choice field it would yield the correct equivalence.

If θ is not an indifference function, equivalence between the expenditures E_t and E_0 are defined by a second-order differential equation (see ECONOMETRICA, January, 1936). This fact can also be expressed by writing the indicator W explicitly in terms of the flexibility θ. First integrate (8) over $d\log E_t$ from some value E_t'. Then take the exponential, notice that (13) $J=dW/d\log E$ and integrate over $d\log E_t$ from some value E_t''. This gives

$$(14)\qquad \int_{E_t''}^{E_t} \exp\left[\int_{E_t'}^{E_t}(1+\theta_t)d\log E_t\right]d\log E_t = \frac{W - W_t''}{J_t'},$$

J_t' and W_t'' denoting the values of J and W in the points E_t' and E_t'' respectively. In other words, along t the double integral in (14) (which is statistically observable in the independence case) *is a linear transformation of the indicator*, the coefficients of the transformation being the special values of J and W indicated. Comparing two paths 0 and t one would, therefore, know the exact equivalence for all expenditure levels if one only knew one pair of points that are equivalent, i.e., have the same W, and one pair that have the same J.

At the morning session of Tuesday the 29th, Mr. D. G. Champernowne (King's College, Cambridge) spoke on "The Theory of Income Distribution." He remarked that if instead of money incomes we take "income power," viz., the logarithm of money income, the observed frequency distributions conform more closely to familiar types of curves. Even the distributions of income powers, however, do not conform closely to the curve of normal error, the Pareto curve, or to any Pearson curve; and Gibrat's modified normal-error curve gives good fits in certain cases only. The contrary impression about the Pareto curve was partly due to the fact that statistics were frequently available for only the rich tail of the frequency curve, where curvature

should in any case be small. It must still be admitted that many observed income-power frequency curves had significantly straight tails when plotted on logarithmic (i.e., Pareto) scale.

For statistics relating to small as well as large incomes the best fit seems to be given by a frequency curve of income power of the type:

$$y = \frac{A}{B \cosh (x - C) - D},$$

where x measures income power,

A is adjusted to give the correct total number of incomes,
B corresponds to the slope of the tail of the Pareto curve,
C is the average income power,
D is adjusted to give the correct kurtosis (or concentration of incomes about the common wage level).

Considering the explanation of different forms, Mr. Champernowne went on to develop the concept of an equilibrium distribution, which may be defined as that distribution for which, in a given social setting, the forces acting to increase inequality are exactly balanced by those tending to decrease it.

In a given society consider all income powers of amount x at the beginning of the year, and let $f(x, y)dy$ be the proportion of those incomes which change during the year between y and $y+dy$.

The cumulants (semi-invariants) $Y_1, Y_2, Y_3, \cdots$ are as usual defined by the identity

$$\log_e \int_{-\infty}^{\infty} e^{Cx} f(x)dx = Y_1 C + \frac{1}{2!} Y_2 C^2 + \frac{1}{3!} Y_3 C^3 + \cdots$$

where $Y_1(x)$ is the mean and $Y_2(x)$ the square of the standard deviation of the frequency distribution $z = f(x, y)$, x being given, and where $Y_3(x)$, $Y_4(x)$ will vanish if, x being given, $z = f(x, y)$ is a normal distribution in y.

The condition for equilibrium is

$$Y_1 + \frac{1}{2!} Y_2 a + \frac{1}{3!} Y_3 a^2 + \cdots = 0,$$

where a is the slope of the income-power frequency curve plotted on logarithmic scale (i.e., of the noncumulative Pareto curve). Y_1, Y_2, etc., will in general be different for different income-power levels x, so the slope of the curve can vary from point to point in equilibrium.

In the case where change of income power from a given level x is the resultant of several independent small causes, $Y_3, Y_4, \cdots$ may be neglected and the condition for equilibrium becomes

$$a = 2Y_1/Y_2.$$

In a country where there was no inheritance, all children had equal opportunities and the Y's were the same for all income powers, the equilibrium distribution of income power would be normal. But if life went on forever, or if each son inherited the same proportion of his father's earning power, there could be no equilibrium, and inequality would increase until some check altered the given conditions.

The second paper of the session was given by M. R. Gibrat (Paris) who spoke on "Depreciation."

Mr. Gibrat said that in many practical cases Kurtz's seven groups of industrial mortality curves (see his book *Life Expectancy of Physical Property*) could be replaced by simpler expressions giving a sufficiently close approximation to the process of depreciation. To derive his formulae Dr. Gibrat took as starting point the fact that a depreciation fund set aside by a corporation is as a rule not placed in stocks or bonds of other corporations but invested in the various assets of the corporation itself. Consequently the *interest rate* on this fund should be determined not as the market rate, but should be derived from the accounts of the corporation itself.

Consider the physical investment activity (building and upkeep of machines, etc.). Part of it is an *initial* investment and part of it is reinvestment caused by an initial investment some time in the past. Consider the year n. Let c_n be the total value of all future reinvestments that are made necessary because of all the initial investments made up to the year n. This total value is taken as discounted down to some conventionally chosen point of time called "today," the computation being made by using "normal" values and a "normal" long-time interest rate j supposed given. Let r_n be the initial investment and t_n the reinvestment respectively made during the year n. We then have

$$c_n = (c_{n-1} + r_{n-1})(1 + j) - t_n.$$

Now let F_n be the depreciation fund as it exists at the end of the year n and let i_n be its (variable) interest rate, which is to be determined. Further let I_n be the rate of earnings during the year n. We then have

$$F_n = F_{n-1}(1 + i_n) + A_n - t_nI_n,$$

A_n being determined by

$$F_n + A_n/j = c_nI_n.$$

By these formulae the computations of i_n and F_n for successive years can be rapidly made using the data contained in the accounts of the corporation.

As a special case the formulae may be used on a number of identical capital goods born at a given instant. The distribution of their deaths may as a rule be represented by a normal curve; and so may the distribution of the deaths in the next generation (only with a larger standard deviation), etc. In this case it can be proved that the correct result is obtained by disregarding the dispersion and figuring as if all the objects died (and were renewed) after exactly the same length of time.

In the afternoon session on Tuesday the 29th, Mr. Horst Mendershausen (University of Geneva) presented "An Example of Meaningful Curvilinear Regressions in Economic Time Series." This paper appears in this issue of ECONOMETRICA.

At the same session Dr. L. Hamburger (Scheveningen) spoke on "The Rise and Decline of Technical Economic Growth Power." He displayed a number of charts in which evidence for the rate of technical innovation, scientific discovery, and commercial expansion, was drawn from statistics of patents, scientific achievement, foreign trade, and population, in different lands and eras. In general he found that these series followed an S-shaped curve, whose maximum rate of growth was about 3 to 4 per cent per annum, and he pointed out that this agreed with the normal rate of long-period interest in nineteenth-century England. He called attention to the similarity of form which his curves displayed, despite the diverse political and social histories of different countries. Taking as his unit of time, t, the working life of an able-bodied man, and adding a factor of decay, he found that the following formula fits the best-known estimates of the growth of the population of the world, Y:

$$Y = 500 + \frac{4500}{4500e^{-0.92t+0.001t^3} + 1}.$$

According to this formula, "world population would have practically attained the upper asymptote belonging to the present secular cycle at the end of the next century." In conclusion, Dr. Hamburger called attention to the social difficulties, through pressure of population, anti-selection, and abuse of technical powers, which attend the saturation phase of the cycle.

The third paper of the session, "Outlines of a Mathematical Theory of Exchange," was given by Dr. Hans Bolza (Würzburg), The purpose of this paper was to isolate the essentials of a theory of catallactics. Considering a case of barter between two group A and B, Dr. Bolza presented a diagram in which the cumulated value of the goods delivered by each group to the other was represented as a function of time, so that the vertical difference between the curve of A's deliveries

and that of B's represented the debt which at a given time one group owed to the other, and the horizontal difference represented the "bartering-time," or lag between the delivery made by one group and the compensating delivery by the other. In conclusion, Dr. Bolza considered the case in which the momentary indebtedness was represented not by promises to deliver specified goods but by generalised means of payment, such as coin, notes, and book credits, in the possession of the creditor, and he examined the condition which must govern the supply of means of payment if the underlying process of barter is not to be distorted by changes in unit purchasing power.

The last paper of the session, entitled "The Measurement of the Sensitivity of Taxes to Fluctuations in Trade," was given by Mr. R. F. Bretherton (Wadham College, Oxford), and has appeared in ECONOMETRICA.[2]

E. H. PHELPS BROWN

New College
Oxford, England

[2] Vol. 5, Apr., 1937, pp. 171–183.

REPORT OF THE DENVER MEETING, JUNE 24–26, 1937

THE AMERICAN summer meeting of the Econometric Society was held in Denver, Colorado on June 24, 25, and 26, 1937, in connection with the one-hundredth meeting of the American Association for the Advancement of Science. Headquarters were at the Cosmopolitan Hotel.

The meeting opened on Thursday morning, June 24, in the auditorium of the Capitol Life Insurance Building, where Dr. Charles F. Roos of the Mercer-Allied Corporation, New York City, presided over a joint session with Section A (Mathematics) and Section K (Economics) of the American Association. The general topic was "Mathematical Economics."

The first paper was by Professor Harold T. Davis of Indiana University and the Cowles Commission for Research in Economics on "The Application of Functionals in Economics." Professor Davis indicated that the origin of this application is to be found in Pareto's theory of *closed and open cycles* published in 1906, a study apparently suggested by Vito Volterra, the father of the modern theory of functionals. The speaker employed the definition: "If a law is given by means of which to every function $x(t)$ defined within an interval $(a \leqq t \leqq b)$, there can be made to correspond one and only one quantity z, then z is said to be a functional of $x(t)$. This relationship is expressed by the symbol $z=F[x(t)]$." Since z depends upon a path of values, line integrals of the type that appear in the theories of utility and indifference surfaces are functionals. Lag correlations and regression lines determined by fitting *open* sets of orthogonal functions to data are other examples of functionals used in economics. The origin of the nonsense correlations observed by Yule and others is frequently to be traced to a failure to recognize this functional nature of regression lines. This difficulty may be avoided by employing functionals in which the kernel is determined from a priori considerations, as in Roos's theory of demand functions of the form

$$y(t) - k = ap(t) + \int_{-\infty}^{t} K(x, t)p(x)dx,$$

rather than by using arbitrary orthogonal sets of functions which determine kernels, $K(x, t)$, that have no economic significance. The paper concluded with a discussion of the functional nature of harmonic analysis. The speaker proved that if $a(T)$ is the Fourier transform, corresponding to a trial period T, of the lag correlation of a series of data, and if $A(T)$ and $B(T)$ are respectively the cosine and sine transforms of the original data, there exists between them the relationship

$$a(T) = A^2(T) + B^2(T).$$

Dr. Gerhard Tintner of the Cowles Commission for Research in Economics gave a paper on "The Economic Meaning of Functionals." He pointed out that much more attention to the use of functionals in economics by the American econometrists has been paid by mathematicians in this field than by professional economists. This can be explained only partly by the highly technical nature of the approach. The true reason is perhaps that they fail to state explicitly the *economic* assumptions and implications of their methods and results. Dynamic problems, like the ones dealt with by the econometrists, are concerned with anticipations and expectations. But the authors in question fail to analyze the mechanism of expectations, which is of primary interest for the economist. This can be exemplified by the dynamic theory of demand, one of the most thoroughly studied subjects. The quantity demanded does not in itself depend on past prices or other past factors. It can be envisaged, on the contrary, to depend on the expected prices and other economic circumstances expected for the future. And these expectations on the other hand will depend largely on past experiences. So, for instance, demand is essentially a functional of future expected prices. Every expected price is a functional of past prices. Since the functional of a functional is again a functional, demand is in the end a functional of past prices, just as the econometrists maintain. But they fail to realize or state explicitly the interconnections outlined here. A thorough study of these interrelationships would involve a consistent theory of choice over an interval of time, or the study of utility as a functional rather than a function.

Mr. Ronald W. Shephard of the University of California spoke on "Incidence of Taxation in a Simplified General Equilibrium." He summarized his paper as follows: "In particular a tax levy of the sales-tax type is introduced into an economic system involving the assumptions of strict competition and a steady state with regard to money. There is implicit in the general categories of the system the further assumption of full employment of resources. a, b, c are distributive coefficients giving the proportion of consumption goods received by producers of capital goods, labor services, and consumption goods respectively. It is found that the changes δa, δb, δc occasioned by the application of the tax are such that $\delta a:\delta b:\delta c = a:b:c$."

The last paper of the session was by Professor Griffith C. Evans and Mr. Kenneth May of the University of California on "The Theory of the Stability of the Regimes of Co-operation and Competition in Business," with reference to simple types of cost tables. With decreasing costs it turns out that a combination like a trust is usually not stable but tends to degenerate either into a quasi monopoly with one or more plants idle, or into some kind of competition. It may happen that

several forms of organization are stable at the same time, and even that the quasi monopoly, which gives the maximum profit to the producer, yields also the maximum benefit to the consumer with a lower price and larger consumption than that of the competitive situation. This is of course not all a universal phenomenon, but depends on particular characters of the cost functions. The theory explains how it happens that businesses manufacturing the same commodity, some with decreasing and some with increasing costs, can compete successfully with each other, and interprets the competitive process by means of which equilibrium is reached, if the final position is stable, and which makes manifest the instability otherwise.

The second session, on Thursday afternoon at the University of Denver School of Commerce, was a joint session with Sections A and K of the American Association and with the American Statistical Association. The subjects was "Mathematical Statistics" and the chair was taken by Professor Edward V. Huntington of Harvard University.

The first paper was by Dr. Francis W. Dresch of the University of California on "Applications of Index Numbers to the Study of General Economic Equilibria," and was a continuation of a paper presented April 3 before the American Mathematical Society at Stanford University. Continuous index numbers of the Divisia type were used to reduce the discussion of general economic equilibria involving a large number of commodities to the discussion of corresponding equilibria for simplified systems involving only a few commodities and a more convenient number of variables. For constant supplies of the primary factors of production, it was found that the index representing the total physical production of consumption goods takes on an extremal value at an equilibrium governed by strict competition. It was shown that for a rather general type of situation, conditions which are sufficient to make this index a maximum are satisfied. As in the earlier paper, the discussion was entirely from a static point of view, and a steady state with respect to the utilization of resources was assumed.

Professor Edward L. Dodd of the University of Texas spoke on "A Certain Index Number as a Mean $f(x_1, x_2, \cdots, x_n)$ with $f(c, c, \cdots, c)$ Defined Only When $c=1$." Professor Dodd summarized his paper as follows: "Suppose that, in a population of males, the proportions in certain age groups which embrace the entire male population are $a, b, \cdots, k$. Then $a+b+\cdots+k=1$. Also, suppose that primed letters indicate the corresponding group proportions for females. Then, likewise, $a'+b'+\cdots+k'=1$. For the first age group, a'/a is the ratio of females to males; for the second age group, b'/b, etc. Let us now make an index number by using any weighted average of these ratios a'/a, b'/b, etc. These ratios are then the arguments or

variates for a mean $f(a'/a, b'/b, \cdots, k'/k)$. But the condition $a'/a = b'/b = \cdots = k'/k$ necessitates that each of these ratios should equal 1 (unity). It turns out, indeed, that $f(1, 1, \cdots, 1)$ equals 1. Hence $f(x_1, x_2, \cdots, x_n)$ is a mean of $x_1, x_2, \cdots, x_n$—whatever these values are—since the only value which the x's can all take on simultaneously is 1, and $f(1, 1, \cdots, 1) = 1$."

The closing paper of the session was by Professor Holbrook Working of the Food Research Institute, Stanford University, on "Relations Among Coefficients of Regression and of Correlation as a Basis for Statistical Inference." Professor Working said that, given the need of measuring the relation between two variates, linearly related, the use of either a coefficient of correlation or a coefficient of regression may appear logically indicated, the choice between these coefficients depending on the aspect of relationship which is to be measured. When a statistical investigation requires analysis of systematic variation among a number of such coefficients derived from a corresponding number of different sets of observations obtained under different conditions, the usual measures of relationship may prove quite unsatisfactory: (1) Changes in the regression coefficients from one set of observations to another, ostensibly reflecting changes in the functional relation between the variates, may actually reflect chiefly changes in the closeness of relationship; or (2) Changes in the correlation coefficients from one set of observations to another, ostensibly reflecting changes in the closeness of relationship between the two variates, may actually reflect in considerable degree changes in the functional relation. This paper presents a method for ascertaining for any particular body of data whether the usual correlation or regression coefficients are satisfactory or suffer from the defects indicated above; and affords a basis, in case the usual coefficients are found unsatisfactory, for selection of more suitable coefficients. Under appropriate conditions, the method affords a practical solution to the problem of measuring the "true" relation between two variates—the relation that would be observed if correlation were perfect and the regression lines coincidental.

On Thursday evening was held a general session with Section K of the American Association, the American Statistical Association, the American Sociological Society, and the Sociological Research Association. Dr. Stuart A. Rice of the Central Statistical Board presided at this session in the Renaissance Room of the University of Denver Library.

Mr. Carl Snyder, for many years with the Federal Reserve Bank of New York, gave an illustrated lecture on "New Foundations for an Economic Science." Mr. Snyder said that it is evident to any one that economics like any other branch of knowledge cannot attain to the dignity of a science until it has a solid foundation in measurements and

correlations. Neither theoretical speculations nor mathematical formulae will supply this unescapable need. Until very recently economics has lacked these measurements and correlations and, as a result, as is clear from the vast variety of opinions and conjectures on almost every economic question, it has simply become lost in the bogs of speculation. Witness some very recent works that vividly illustrate the situation: Hundreds of pages of theory or algebraic equations, with almost no support from measurements and significant facts. We have now, however, the widest variety of compilations and indexes, and the beginnings of some extremely interesting and valuable correlations. So we now know, e.g., that an index of commodity prices at wholesale is not an adequate measure of the general price level or average of all kinds of prices included in payments. We have at least an approach to such a general price level for both the United States and Great Britain. This, with new measures of production and trade, and of the velocity or rate of turnover of bank deposits, has made possible a rational or quantitative measure of the relation of money, credit, and prices, so vital for the present economic situation. And so with a wide variety of other measures, calculations, and comparisons. With the aid of all these, and their ranges wide, it is now possible to test almost every variety of economic theory or hypothesis and to establish on a solid quantitative basis their truth or falsity. In another generation it is probable that economists will wonder why such quantitative measures should not have been undertaken long ago. As a matter of fact their beginnings date back to the work of Jevons in 1863. Progress has been slow but the accumulation is already large.

The fourth session, on Friday morning, June 25 at the University of Denver School of Commerce, was a joint session with Section K of the American Association, the American Sociological Society, and the Sociological Research Association. The general topic was "Economic Balance and Impacts," and the chair was taken by Professor Griffith C. Evans of the University of California.

The first speaker was Dr. Charles F. Roos of the Mercer-Allied Corporation, who spoke on "Measuring Economic Impacts." Dr. Roos discussed the effects of selected economic impacts upon the number of housing units built, employment, production, and stock prices. By reference to various correlations which had held over considerable periods of time, he urged that it is pointless to argue what the effects of impacts were until the relative influences of fundamental factors were known. For instance, the effects of the Federal Housing Administration Act could be measured only after effects of changes in rent, occupancy cost, and normal credit conditions were first measured. Effects of the NRA on employment and production could be measured

only in conjunction with measures of changes in monetary supply, monetary velocity, and federal spending. He then showed how this could be accomplished through selection of industries and by relation of their statistics to general factors. In discussing the effects of economic impacts on the stock market he said that very often when news is announced, which according to economic theory should send stock prices higher, they actually went lower or vice versa. He explained this phenomenon by pointing out that the particular effect obtained depended upon the position of stock prices in their ranges as determined by earnings records and outlooks. Also of extreme importance in pricing stocks according to news items is the inertia or momentum of prices during the preceding few days or weeks.

The second paper was by Mr. Carl Snyder on "High Profits and High Wages," and was illustrated by lantern slides. Mr. Snyder gave a brief summary of an extended research into the mechanism by which high wages and general well-being in a progressive country are attained. It seems now conclusively demonstrated that high wages are derived solely from the increase of the product per worker and that this increase in wages is automatic and cannot be checked or promoted by employers, or labor unions, or legislation. In turn this increase in the product per worker is due, for all practical purposes, solely to the increase of the mechanical power employed and new inventions, devices, and processes. The so-called increase in "efficiency" of the workers is entirely a myth. In turn this increase in mechanical power requires a corresponding increase in the capital investment, and this in almost exact ratios; that is to say, for every increase of a million or a billion dollars in the value of manufactures per annum there has been required the investment of a million or billion dollars of new capital. It is, therefore, through this investment of capital, solely, that higher pay and shorter hours in industry have been possible. In turn it is now established that this capital is derived in a very large part either from the savings of the corporations and large enterprises, directly, in the form of a surplus; or from the very small number of individuals who own shares in these enterprises. The proof is that control of domestic corporations doing 80 or 90 per cent of the nation's business is in the hands of perhaps less than a hundred thousand people, and this, so far from being an evil, brings all the advantages of the highest efficiency and the greatest possible capital savings.

The last paper of the morning was by Mr. Louis H. Bean of the Agricultural Adjustment Administration, U. S. Department of Agriculture, on "The Balance Between Agriculture and Industry." Mr. Bean summarized his paper as follows: "While this subject can be discussed from numerous angles, this paper deals chiefly with the changing

long-time relationships between agricultural and nonagricultural production, prices, and incomes. A knowledge of the basic factors in these relationships is necessary in appraising current agricultural policy as implied in the phrases "parity prices" and "parity income," and in the "ever-normal granary" idea behind efforts for more regular agricultural output. The outstanding factors in both price and income relationships are the relative rates of production in agriculture and industry. The greater rise in industrial than in agricultural production per capita over the past century has given an increased purchasing power to agricultural prices. The recent changes in the relation of income per capita in agriculture to income per capita for the rest of the population can be explained statistically in terms of three factors, industrial output, agricultural output, and foreign demand, the relative importance of which appear to have been 5, 3, and 2 respectively. A return to prewar parity prices and income, in view of increasing farm production and the improvement in foreign demand, hinges on a very substantial rise in industrial activity, this being the source of the real national income and the domestic demand for farm products."

The final session, on "Inflation," was held on Saturday morning, June 26, at the University of Denver School of Commerce. Mr. Carl Snyder presided.

The first paper was by Professor E. J. Working of the University of Illinois on "The Course of Prices of Agricultural Commodities and their Outlook." Professor Working stated that in analyzing factors which have been affecting the average level of agricultural prices it was highly important to consider the agricultural situation as a whole, because substitutes for agricultural products as a whole are very limited. Hence the elasticity of demand for the whole is much larger than the average elasticity of demand for the various individual commodities. An analysis of changes during the past fifteen years in an index of prices of agricultural products indicated that during the period the effect of demand changes was twice as great as the effect of changes in supply. For the coming year, however, it is probable that changes in supply will be more important than changes in demand, because carry-over into 1937 will be the smallest in eight years and present prospects are that production for 1937 will be the largest in five years. There is fairly strong evidence that we may expect an increase in supplies sufficient to reduce prices considerably, unless there should be a counterbalancing increase in demand. Changes in domestic demand are quite satisfactorily measured by changes in the nonagricultural income of the people of the United States. Changes in foreign demand are due to a variety of factors, and may be measured roughly by the value of our agricultural exports. There is considerable evidence

to indicate that price-level movements are ordinarily the result of changes in the value of "flexible price" commodities like agricultural products. Hence prospects for an increased supply must be considered as an important possible cause of a minor decline in the average level of wholesale commodity prices during the next two years.

The second paper was by Professor James Harvey Rogers of Yale University on "Some Quantitative Aspects of Inflation." Professor Rogers, for the purposes of this paper, defined inflation as an increase in the purchasing medium in the hands of purchasers without corresponding increase in goods to be purchased. He presented and described at length a balance sheet of the monetary system of the United States, which gave a quantitative statement of the monetary base and its utilization. Formulae were derived for maximum loan and deposit expansion of the American banking system under various hypotheses, and numerical computations of important inflation possibilities were presented. It was concluded that the most serious threat of inflation in the United States comes from big Treasury deficits, and that all purely monetary influences leading to inflation could be controlled, unless Treasury financing should continue very large. The following consequences of the recent great increase in the monetary gold stock of the world were pointed out: (1) In the long run either the general level of prices must rise considerably or else the price of gold must be lowered; the only other possibilities being the demonetization of gold or the reduction of our protective tariff. (2) Any substantial reduction in the dollar price of gold would probably be accompanied by a similar, though slightly smaller, reduction in its sterling price. Hence the impact on the prices of agricultural products and on those of other articles with markets chiefly in Great Britain or the Sterling Group would at least be mitigated. (3) The major effects of a lowering of the gold price —however unlikely its use—could be accomplished by merely reducing sufficiently the *buying price* of *imported* gold supplies, without changing the price of $35 an ounce for the purchase of domestic supplies and for all sales, foreign and domestic. In this way, further embarrassing gold imports could be avoided while foreigners could be prevented temporarily and perhaps permanently from taking a profit on their recent sales of gold to us or from withdrawing it too suddenly for our convenience. Such a move, however, would mean a new departure from the gold standard and the dollar could be expected to rise in the exchange markets of the world and to fluctuate within wide limits.

The last paper was by Professor Jesse H. Bond of the University of Oregon on "A Consumptionstat: Its Necessity as an Inflation Control." Professor Bond said that machine technology, under present accounting and credit methods, creates deficits and surpluses of con-

sumer income. Deficits make inflationary policies inevitable. The consumptionstat is a proposed mechanism for applying the surpluses against the deficits and thus keeping the total money volume of consumptive demand reasonably in step with the total physical volume of production. The consumptionstat will use the commodity price level as a barometer of consumptive demand. An all-commodity wholesale price index, adjusted for decreasing costs due to technical progress, will be adapted to read 100 at a chosen normal price level. A sales-tax administrative set-up will do two things: first, capture price-raising surpluses of consumer money income by imposing a sales tax when the price-level index crosses 105, say, going up, and collecting it until the index again becomes 100; second, remove deficits in consumer purchasing power by starting to issue official federal trading stamps when the commodity price-level index crosses 95 going down, issuing and redeeming them in cash until the index returns to normal. Government credit will be used as a necessary cushion, the public debt rising as trading stamps are issued, being retired as the sales tax is collected.

DICKSON H. LEAVENS

Cowles Commission for
Research in Economics
Colorado Springs, Colorado

ANNOUNCEMENT OF THE ATLANTIC CITY AND INDIANAPOLIS MEETINGS, DECEMBER 27–30, 1937

THE ECONOMETRIC SOCIETY will hold meetings in Atlantic City Monday and Tuesday, December 27 and 28, 1937 and in Indianapolis Wednesday and Thursday, December 29 and 30. One session in Atlantic City will be a joint session with the American Statistical Association and one session in Indianapolis will be joint with the Institute of Mathematical Statistics and Section A (Mathematics) of the American Association for the Advancement of Science. Among those who have indicated willingness to speak at these meetings are: Harold T. Davis, Alvin H. Hansen, Harold Hotelling, P. Luzzato-Fegiz, Francis McIntyre, James Harvey Rogers, Charles F. Roos, Gerhard Tintner, Victor S. von Szeliski, and Roswell H. Whitman. Sessions will be devoted to monetary theory, national income, theory of employment and wages, price analysis, welfare economics, statistical techniques, and miscellaneous papers. All members of the society who are working on research which they expect to complete by the time of the meetings are urged to write immediately to the chairman of the program committee, Dr. Charles F. Roos, 420 Lexington Avenue, New York City.

LIST OF MEMBERS OF THE ECONOMETRIC SOCIETY

As of October 1, 1937.

Please notify the Secretary, Alfred Cowles 3rd, 301 Mining Exchange Building, Colorado Springs, Colorado, U. S. A., of any changes of address or of any errors in names, titles, or addresses as printed.

ADAMSON, MR. WENDELL, Statistician, University of Alabama, University, Alabama.

AKERMAN, PROF. GUSTAV, 2 Södra Väjen, Göteborg, Sweden.

AKERMAN, DR. JOHAN, University of Lund, Lund, Sweden.

ALLEN, MR. R. G. D., London School of Economics, Houghton Street, Aldwych, London, W. C. 2., England.

ALLIX, PROFESSEUR EDGARD, Faculté de Droit, l'Université de Paris, 7 rue Ribéra, Paris XVI, France.

ALSBERG, DR. CARL L., Food Research Institute, Stanford University, Palo Alto, California.

ALTSCHUL, DR. EUGEN, School of Business Administration, University of Minnesota, Minneapolis, Minnesota.

AMOROSO, PROF. LUIGI, Università di Roma, Roma, Italy.

ANDERSON, PROF. M. D., University of Florida, Gainesville, Florida.

ANDERSON, PROF. OSKAR N., Director of the Statistical Institute for Economic Research, State University, Shipka 6, Sofia IV, Bulgaria.

ANDREW, MR. SEYMOUR L., American Tel. & Tel. Co., 195 Broadway, New York City.

ANGELL, PROF. JAMES W., Columbia University, New York City.

ANTONELLI, PROF. ÉTIENNE, Professeur de Faculté, 173 rue de Courcelles, Paris, XVII, France.

AOYAMA, MR. HIDEO, Faculty of Economics, Kyoto Imperial University, Kyoto, Japan.

ARON, DR. ARNO, Schumannstrasse 20, Frankfurt (Main), Germany.

AUPETIT, DR. ALBERT, 2 avenue Montaigne, Paris VIII, France.

AXE, MR. E. W., 730 Fifth Avenue, New York City.

AYRES, COL. LEONARD P., Cleveland Trust Co., Cleveland, Ohio.

BACHI, PROF. RICCARDO, Via Antonio Bosio N. 2, Roma, Italy.

BAGGE, PROF. GOSTA, Stockholms Högskola, Stockholm, Sweden.

BAINOFF, DR. IVAN, 28 rue Graf Ignatieff, Sofia, Bulgaria.

BAKKER, DR. O., Dedelstraat 18, 's Gravenhage, The Netherlands.

BALÁS, DR. KÁROLY, Kecskeméti-utca 9, Budapest, Hungary.

BALKRISHNA, PRINCIPAL, Rajaram College, Kolhapur, India.

BANOS, PROF. OLEGARIO, Banca de España, Madrid, Spain.

BARBIERI, MR. GINO, Università Cattolica del S. Cuore, Piazza S. Ambrogio 9, Milano, Italy.

BARDET, M. GÉRARD, 145 avenue du Général Michel Bizot, Paris XII, France.

BARFOD, MR. BØRGE, Valentinersvej 30, København, F., Denmark.

BARRIOL, PROF. ALFRED, 27 rue Mogador, Paris IX, France.

BASU, PROF. PRAPHULLACHANDRA, Holkar College, Indore, C. I., India

BATHGATE, MR. ARTHUR E., Teulon, Manitoba, Canada.

BAYKITCH, M. VELIMIR, 59 rue Sanje Zivanovica, Beograd, Yugoslavia.

BELCHER, MR. DONALD R., Asst. Chief Statistician, American Tel. and Tel. Co., 195 Broadway, New York City.

BELSCHNER, MISS MARTHA C., 1342 N. El Paso St., Colorado Springs, Colorado.

BELZ, MR. M. H., The University of Melbourne, Carlton N. 3, Victoria, Australia.

BENTLEY, DR. A. F., Paoli, Indiana.

BERNARDELLI, MR. H., 20 Pagan Road, University Estate, Rangoon, Burma, India.

BESSO, ING. M. A., Bureau Fédéral de la propriété intellectuelle, Bern, Switzerland.

BEVERIDGE, SIR WILLIAM H., London School of Economics, Houghton Street, Aldwych, London, W. C. 2, England.

BICKERDIKE, MR. C. F., 12 Hengrave Road, Forest Hill, London, S. E. 23, England.

BIJL, MR. A., Rivierenlaan 72I, Amsterdam Z, The Netherlands.

BILIMOVIC, PROF. AL., Univerza Aleksandra, Ljubljana, Yugoslavia.

BISSELL, MR. RICHARD M., 1292 Yale Station, New Haven, Connecticut.

BJORAA, MR. S. J. "Andvake," 30 Stortings gt, Oslo, Norway.

BLAIR, MR. WILLIAM, Balliol College, Oxford, England.

BLAKSTAD, MR. EGIL, Chief Engineer, De No Fa, Fredrikstad, Norway.

BODE, DR. KARL, St. John's College, Cambridge, England.

BOEHM, DR. CARL, Gustloffplatz 2, Berlin-Charlottenburg 5, Germany.

BOGART, PROF. ERNEST L., 1101 W. California Street, Urbana, Illinois.

BOLZA, DR. H., Postamt 7, Würzburg, Germany.

BOLZINGER, ING. A., 16 quai des Tanneurs, Nantes (Loire-Inférieure), France.

BONINSEGNI, PROF. P., Université de Lausanne, Lausanne, Switzerland.

BORDIN, PROF. ARRIGO, V. Carita 32, Catania, Sicily, Italy.

BOREL, PROF. ÉMILE, Faculté des Sciences, Université de Paris, Paris, France.

BORGATTA, PROF. GINO, N. 1, Piazzale Baracco, Milano, Italy.

BOUNIATIAN, DR. MENTOR, 97 rue Lauriston, Paris XVI, France.

BOUSQUET, PROF. G. H., Université d'Alger, Alger, Algeria.

BOWERMAN, MR. WALTER G., Asst. Actuary, New York Life Ins. Co., New York City.

BOWLEY, PROF. A. L., Marley Hill, Haslemere, England.

BRANNON, DR. MAXIMILIANO PATRICIO, San Salvador, El Salvador, Central America.

BRATT, PROF. E. C., Lehigh University, Bethlehem, Pennsylvania.

BREIT, MR. MAREK, Polish Institute for Economic Research, Elektoralna 2, Warszawa, Poland.

BRESCIANI-TURRONI, PROF. DR. COSTANTINO, Faculty of Law, Giza, Egypt.

BRITZELMAYR, DR. W., Ludwigstrasse, 9, München, Germany.

BROWN, MR. ARTHUR JOSEPH, Queen's College, Oxford, England.

BROWN, MR. E. H. PHELPS, New College, Oxford, England.

BROWN, PROF. THEODORE HENRY, Harvard University, Cambridge, Mass.

BRULÉ, M. PAUL E., 161 rue Berkendael, Bruxelles, Belgium.

BUFANO, DR. GIUSEPPE, Viale Gorizia, 43, Roma, Italy.

BURGESS, DR. ROBERT W., 440 Pelham Manor Road, Pelham Manor, New York.

BURKHARDT, DR. FELIX, Alemannenstrasse 20, Dresden, Germany.

BUSHEY, DR. J. HOBART, Hunter College of the City of New York, Park Avenue and 68th St., New York City.

BYÉ, PROF. MAURICE, Faculté de Droit, 15 rue de Languedoc, Toulouse, France.

CAIRNCROSS, MR. A. K., Department of Political Economy, University, Glasgow, Scotland.

CAMPBELL, MR. C. D., School of Social Sciences and Administration, Department of Commerce, 19 Abercromby Square, Liverpool, England.

CANNING, PROF. JOHN B., Stanford University, Palo Alto, Calif.

CARMILLE, M. RENÉ, 30 avenue Brimborion, Sèvres (Seine et Oise), France.

CARVER, PROF. HARRY C., University of Michigan, Ann Arbor, Michigan.

CHAIT, MR. B., 38 rue Gysels, Anvers, Belgium.

CHAMPERNOWNE, MR. D. C., King's College, Cambridge, England.

CHANG, MR. C. C., Bank of China, Shanghai, China.

CHAO, DR. J. T., Professor of Economics, Department of Economics, Tsing-Hua University, Peiping, China.

CHAYROU, INTENDANT GÉNÉRAL, 6 rue Rosa Bonheur, Paris XV, France.

CHEN, PROF. P. C., Director, Bureau of Statistics, Ministry of Industry, Nanking, China.

CHEN, DR. WARREN, Directorate of Statistics, National Government, Nanking, China.

CHERINSKY, MR. ELMER, Aberdeen Proving Ground, Maryland.

CHRISTENSON, PROF. C. L., Department of Economics, Indiana University, Bloomington, Indiana.

CLARK, MR. COLIN, 31 Parkside, Cambridge, England.

CLARK, PROF. J. M., Columbia University, New York City.

CLARK, DR. VICTOR S., Consultant in Economics, Library of Congress, Washington, D. C.

COATS, MR. R. H., Dominion Statistician, Dominion Bureau of Statistics, Ottawa, Ontario, Canada.

COENEN, DR. P. A., C. Reinierszkade 25, 's Gravenhage, The Netherlands.

COHEN, MISS RUTH, Agricultural Economics Research Institute, Oxford, England.

COLM, PROF. DR. G., 100 Buckingham Rd., Yonkers-Nepperhan, N. Y.

COLSON, PROF. C., Membre de l'Institut, 2 rue de la Planche, Paris VII, France.

COMPTON, DR. WILSON, 2900 Cathedral Avenue, Washington, D. C.

COPELAND, DR. MORRIS A., Central Statistical Board, 7028 Commerce Bldg., Washington, D. C.

COURT, MR. ANDREW TRAWICK, Automobile Manufacturers Association, General Motors Bldg., Detroit, Michigan.

COVER, PROF. JOHN H., Faculty Exchange, Box 216, University of Chicago, Chicago, Illinois.

COWLES, MR. ALFRED, 3RD, Mining Exchange Building, Colorado Springs, Colorado.

CRACCO, PROF. FRANÇOIS, 32 boulevard de Erlemont, Louvain, Belgium.

CRAIG, MR. JAMES IRELAND, Financial Secretary, Ministry of Finance, Cairo, Egypt.

CRAMER, PROF. HARALD, Skarviksv. 7, Djurholm, Sweden.

CRATHORNE, DR. A. R., University of Illinois, Urbana, Illinois.

CREEDY, PROF. F., 307 Hillsboro Street, Chapel Hill, N. Carolina.

CROXTON, PROF. FREDERICK E., Dept. of Statistics, Columbia University, New York City.

DALMULDER, DR. J. J. J., Westzeedijk 76, Rotterdam, The Netherlands.

D'ANNA, DR. FRANCESCO COPPOLA, Confederazione Generale Fascista delle Industria Italiana, Piazza Venezia, 11, Roma, Italy.

DANSON, MR. FORREST M., Cowles Commission for Research in Economics, Mining Exchange Building, Colorado Springs, Colorado.

DARLING, DR. GEORGE B., W. K. Kellogg Foundation, Battle Creek, Michigan.

DARMOIS, PROF. G., 7 rue de l'Odéon, Paris VI, France.

DARRIEUS, ING. M. G., 7 avenue du Général Maunoury, à Houilles (Seine-et-Oise), France.

DASCOVICI, PROF. N., l'Université de Jassy, Iasi, Rumania.
DAVENPORT, PROF. DONALD H., Graduate School of Business Administration, Harvard University, Boston, Mass.
DAVID, MISS FLORENCE N., Department of Statistics, University College, London, W.C.1., England.
DAVIES, PROF. G. R., College of Commerce, University of Iowa, Iowa City, Iowa.
DAVIES, MR. THURSTON J., President, Colorado College, Colorado Springs, Colo.
DAVIS, PROF. H. T., Dept. of Mathematics, Northwestern University, Evanston, Illinois.
DE ARTIGAS, PROF. JOSÉ ANTONIO, c/o The National City Bank of New York, 60 avenue des Champs Élysées, Paris VIII, France.
DE BOÉR, DR. ELEK, Department of Law, Kecskemét, Hungary.
DE BORDES, DR. J. VAN WALRÉ, 4 Chemin du Pré de l'Ours, Genève, Switzerland.
DE FELLNER, DR. FREDERIC, Palatin Joseph University, Budapest, Hungary.
DE LANGEN, MR. W. J., Duinvoetlaan 5, Wassenaar, The Netherlands.
DELLA-RICCIA, DR. ANGELO, 253 Chaussée d'Alsemberg, Bruxelles, Belgium.
DEL VECCHIO, PROF. GUSTAVO, R. Università, Via delle Lame 8-10, Bologna, Italy.
DEMARIA, PROF. GIOVANNI, Università Commerciale Luigi Bocconi, Milano, Italy.
DE PIETRI-TONELLI, PROF. ALFONSO, R. Istituto superiore di scienze economiche, s. Marta 21-33, Venezia, Italy.
DE PONTEVES, PROFESSEUR, 31 parc de Montretout, Saint Cloud (Seine et Oise), France.
DERKSEN, DR. J. B. D., Sportlaan 81, 's Gravenhage, The Netherlands.
DESSAU, MR. EINAR, Director Tuborg Brewery, Hellerup, Denmark.
DE SZADECZKY-KARDOSS, PROFESSOR TIBOR, University, Debrecen, Hungary.
DE WOLFF, DRS. P., Boksdoornstraat, 23, 's Gravenhage, The Netherlands.
DE WOLFF, DR. S., Rubensstraat, 93, Amsterdam (Z), The Netherlands.
d'HARCOURT, M. JACQUES, 71 faubourg Saint-Honoré, Paris VIII, France.
DICKENS, MR. ALBERT E., The Stevens Apts., 720 S. Michigan Ave., Chicago, Illinois.
DICKINSON, PROFESSOR FRANK G., 411 Commerce Building, University of Illinois, Urbana, Illinois.
DIEULEFAIT, PROF. CARLOS E., Professor de Estadistica, Universidad Litoral, Rosario, Argentina.
DIVISIA, PROF. FRANÇOIS, 95 boulevard Jourdan, Paris, XIV, France.
DODD, PROF. EDWARD L., 3012 West Avenue, Austin, Texas.
DODGE, MR. HAROLD F., Bell Telephone Laboratories, 463 West St., New York City.
DOMINEDO, DR. VALENTINO, Università Luigi Bocconi, Milano, Italy.
DONHAM, DEAN WALLACE B., Graduate School of Business Administration, Harvard University, Soldiers Field, Boston, Mass.
DONOVAN, COLONEL WILLIAM J., 2 Wall Street, New York City.
DOSTAL, PROF. B. F., Mathematics Department, University of Florida, Gainesville, Florida.
DOUGLAS, PROF. PAUL H., University of Chicago, Chicago, Illinois.
DUBRUL, MR. STEPHEN M., General Motors Corporation, Detroit, Michigan.
DUCASSÉ, M. F. A., 127 rue du Marche, Bruxelles, Belgium.
DUMONT, ING. AUGUSTIN, 80 rue de Monceau, Paris VIII, France.
DUNCAN, MR. A. J., 137 Jefferson Road, Princeton, New Jersey.
DUNNIGAN, MR. WILLIAM, 111 No. Wacker Drive, Chicago, Illinois.
DUPRIEZ, PROF. LÉON H., 145 rue des Moutons, Louvain, Belgium.
DUTILH, MR. J. R., 46 boulevard St. Michel, Paris, France.

EDIE, DR. LIONEL D., Pres. Capital Research Co., Inc., 20 Exchange Place, New York City.

EDWARDS, MR. EDWARD, School of Business Administration, Indiana University, Bloomington, Indiana.

EDWARDS, PROF. RAMÓN SALAS, Dieciocho 277, Santiago, Chile.

EINAUDI, PROF. LUIGI, via Lamarmora, 60, Torino, Italy.

EISENHART, MR. C., Department of Mathematics, University of Wisconsin, Madison, Wisconsin.

ELFVING, DR. GUSTAV, Abo Akademie, Turku, Finland.

ELSAS, DR. MORITZ JOHN, 29 Steeles Road, London, N.W. 3, England.

ENGLIS, PROF. KAREL, Governor, National Bank of Czechoslovakia, Praha, Czechoslovakia.

EVANS, PROF. G. C., University of California, Berkeley, California.

EZEKIEL, DR. MORDECAI, Secretary's Office, Department of Agriculture, Washington, D. C.

FANNO, PROF. MARCO, Università di Padova, Padova, Italy.

FELLNER, DR. WILLIAM, Benczur utca II, Budapest, VI, Hungary.

FENCHEL, DR. WERNER, Matematisk Institut, Blegdamsvej, 15, København, Denmark.

FERRARA, DR. COSTANZA, Istituto Nazionale delle Assicurazioni, Roma, Italy.

FERRARI, PROF. GIOVANNI, Scuola di Statistica, Regia Università di Padova, Padova, Italy.

FISHER, MR. ERNEST M., Federal Housing Administration, Washington, D. C.

FISHER, PROF. IRVING, Yale University, New Haven, Conn.

FLANDERS, MR. RALPH E., Box 477, Springfield, Vermont.

FLUX, SIR ALFRED, Kronprinsvej 12, København, Denmark.

FOGELSON, MR. S., Office Central de Statistique, Aleje Jerozalimskie 32, Warszawa, Poland.

FONTIGNY, MR. PAUL, Zolder, Limburg, Belgium.

FORTUNATI, PROF. PAOLO, Scuola di Statistica, R. Università di Padova, Padova, Italy.

FRANCK, DR. L. ROSENSTOCK, 55 rue de Lille, Paris, France.

FRANKEL, PROF. S. HERBERT, University of Johannesburg, Johannesburg, S. Africa.

FRAZER, MR. GEORGE E., Board of Trade Building, 39th Floor, Chicago, Illinois.

FRÉCHET, PROF. MAURICE, Faculté des Sciences de l'Université de Paris, 11 rue Pierre-Curie, Paris V°, France.

FREEMAN, MR. H. A., Instructor in Economics and Statistics, Massachusetts Institue of Technology, Cambridge, Massachusetts.

FRERICHS, ING. CHARLES, 66 rue des Colonies, Bruxelles, Belgium.

FRICKEY, DR. EDWIN, 44A Holyoke House, Harvard Square, Cambridge, Mass.

FRIDAY, DR. DAVID, 4525 Garfield Street, Washington, D. C.

FRIEDRICH, DR. W. G., Jungmannovo nam. 8a, Praha-Karlin, Czechoslovakia.

FRISCH, PROF. RAGNAR, Slemdalsveien 98, Oslo, Norway.

FRUMENTO, DR. ARMANDO, dell' Universitá Commerciale Luigi Bocconi, Viale Lazio 16, Milano, Italy.

FRY, DR. THORNTON C., Bell Telephone Laboratories, 463 West Street, New York City.

FURLAN, DR. LOUIS-VLADIMIR, Professor at the University, Basle, Switzerland.

FUSI, DR. FRANCO, Via Petrarca 15, Milano, Italy.

GAITSKELL, MR. HUGH, University College, Gower Street, London, W.C.I. England.

GALLIOT, M. A., 10 rue César Franck, Paris XV, France.

GARINO CANINA, PROF. ATTILIO, Corso Francia 100, Torino, Italy.

GARVER, PROF. FREDERIC B., University of Minnesota, Minneapolis, Minnesota.

GELDENHUYS, DR. F. E., Commercial Counsellor, South African Legation, Piazza dell' Independenza, No. 3, Roma, Italy.

GEORGESCU, DR. NICOLAS, 7-bis Pescari, Bucuresti IV, Rumania.

GHEORGHIU, MR. DIM. I., Boulevard Vintilla, Brătianu 47, Bucuresti, Rumania.

GIBLIN, PROF. L. F., University, Melbourne N. 3, Australia.

GIBRAT, DR., 10 rue Oswaldo Cruz, Paris XVI, France.

GIFFORD, MR. C. H. P., 32 Stafford Street, Edinburgh 3, Scotland.

GILBOY, DR. ELIZABETH W., Committee on Research in the Social Sciences, Harvard University, Cambridge, Mass.

GINI, PROF. CORRADO, Istituto di Statistica e Politica Economica, R. Università, Via delle Terme di Diocleziano, 10, Roma, Italy.

GIVENS, DR. MEREDITH B., Room 1127, 342 Madison Ave., New York, N. Y.

GLOERFELT-TARP, MR. B., Supervisor, Great Northern Telegraph Co., København, Denmark.

GOBBI, PROF. ULISSE, Università Commerciale Luigi Bocconi, Milano, Italy.

GOLDSCHMIDT, DR. R. W. c/o Securities and Exchange Commission, Analytical and Technical Division, 1778 Pennsylvania Ave., Washington, D. C.

GOLICK, MR. A. FRANK, 30 North Curtis Street, Chicago, Illinois.

GOLOVIN, MR. NICHOLAS, c/o R. H. Macy & Co., 34th and Broadway, New York City.

GOUDRIAAN, PROF. J., JR., High School of Commerce of Rotterdam, Einhoven, The Netherlands.

GRAFÉ, ING. PAUL, 32 avenue de l'Indépendance Belge, Koekelberg-Bruxelles, Belgium.

GRAFF, PROF. F. W., John Carroll University, Cleveland, Ohio.

GRANT, PROF. EUGENE L., 2045 Tasso Street, Palo Alto, California.

GRAZIANI, PROF. AUGUSTO, via Andrea D'Isernia, 4, Napoli, Italy.

GREBE, DR. JOHN J., 612 East Pine, Midland, Michigan.

GRIFFIN, PROF. F. L., Department of Mathematics, Reed College, Portland, Oregon.

GROVE, DR. CHARLES C., 143 Milburn Avenue, Baldwin, Long Island, New York.

GUARESTI, DR. JUAN JOSÉ, Paseo Colón 185, Buenos Aires, Argentina.

GUMBEL, PROF. E. J., 1 rue P. Huvelin, Lyon, France.

GUTHMAN, PROFESSOR HARRY G., Northwestern University, Commerce Building, Evanston, Ill.

HAAVELMO, MR. TRYGVE, The University Institute of Economics, Oslo, Norway.

HABERLER, DR. GOTTFRIED, Department of Economics, Harvard University, Cambridge, Mass.

HADAMARD, PROF. JACQUES, l'École Polytechnique, 25 rue Jean Dolent, Paris VII°, France.

HAGSTROEM, DR. K.-G., Actuary, "Framtiden," Stockholm 3, Sweden.

HALM, PROF. GEORG, Tufts College, Massachusetts.

HAMBURGER, DR. IR. L., Leuvenschestraat 108, Scheveningen, The Netherlands.

HAMILTON, PROF. THOMAS R., Agriculture Building, College Station, Texas.

HAMMARSKJÖLD, DR. DAG., Sveriges Riksbank, Stockholm, Sweden.

HANAU, DR. ARTHUR, Ithweg 8, Berlin-Zehlendorf, Germany.

HANSEN, PROF. ALVIN H., Graduate School of Public Administration, Harvard University, Cambridge, Mass.

HANSEN, MR. A. S., Consulting Actuary, Board of Trade Building, Chicago, Ill.

HARROD, MR. R. F., Christ Church, Oxford, England.

HARTKEMEIER, PROF. HARRY PELLE, 116 Business and Public Administration Building, University of Missouri, Columbia, Missouri.

HAWTREY, MR. R. G., 18 Elm Park Gardens, London, S. W., England.

HAYAKAWA, PROF. M., Faculty of Agriculture, Imperial University, Sapporo, Japan.

HAYEK, PROF. DR. FRIEDRICH A., 15 Turner Close, London, N. W. 11, England.

HAYES, DR. ANSON, Director of Research, American Rolling Mills, Middletown, Ohio.

HEILPERIN, DR. M. A., c/o Institut Universitaire de Hautes Études Internationales, 5 promenade du Pin, Geneva, Switzerland.

HELLER, PROF. DR. WOLFGANG, Müegyetem-Rahpart 3, Budapest, Hungary.

HENDERSON, MR. A., King's College, Cambridge, England.

HENDERSON, DR. ROBERT, Crown Point, New York.

HERMANT, M. MAX, 10 rue Alfred Roll, Paris XVII, France.

HERVEY, DR. GEORGE W., Consumers' Counsel Division, U. S. Dept. of Agriculture, Washington, D. C.

HEYDEL, PROF. ADAM, Jagellonian University, Krakow, Poland.

HICKS, MR. J. R., Gonville & Caius College, Cambridge, England.

HILL, MR. DANIEL A., 12503 Lancelot Avenue, Cleveland, Ohio.

HIRSHFELD, MR. C. F., Detroit Edison Co., Detroit, Michigan.

HITCH, MR. CHARLES J., Queen's College, Oxford, England.

HO, DR. FRANKLIN L., Institute of Economics, Nankai University, Tientsin, China.

HOOK, MR. JAMES W., Geometric Tool Company, New Haven, Connecticut.

HORST, DR. PAUL, Procter & Gamble Company, Cincinnati, Ohio.

HOTELLING, PROF. HAROLD, Columbia University, New York City.

HRASE, DR. JOSEF, Ovenecka, 23B, Praha XIX, Czechoslavakia.

HUBER, DR. M., Director d'Études de l'Institut de Statistique, de l'Université de Paris, 173 boulevard Saint-Germain, Paris, France.

HUTCHINS, MR. JOHN G. B., 16 Stearns Road, Watertown, Massachusetts.

HYTTEN, PROF. TORLIEV, Head Office, Bank of New South Wales, Box 2722 C. G. P. O., Sydney, New South Wales.

INSOLERA, PROF. FILADELFO, Via Velletri, 19, Roma, Italy.

IONESCU SISESTI, PROF. G., l'Académie des Hautes Études Agricoles, Bucuresti, Rumania.

ISRAEL, M. MAURICE, Ingenieur en Chef des Manufactures de l'État, 15 boulevard Lannes, Paris XVI, France.

IVERSEN, DR. CARL, Docent Københavns Universitet, København, Denmark.

JACOBSON, MISS KATHERINE, Apt. 610, 1722 19th Street, N.W., Washington, D. C.

JACOBSSON, MR. PER., Economic Adviser to the Bank for International Settlements, Basle, Switzerland.

JAFFÉ, PROF. WILLIAM, Department of Economics, Northwestern University, Evanston, Illinois.

JAHN, DIRECTOR GUNNAR, Det Statistiske Centralbyra, Oslo, Norway.

JAIN, DR. L. C., Professor of Economics, The University, Lahore, India.

JAMES, MR. WENDELL A., Farm Credit Administration, Washington, D.C.

JANKO, DR. JAROSLAV, Kolinska 12, Praha XII, Czechoslovakia.

JANNACCONE, PROF. PASQUALE, 107 Corso Galileo Ferraris, Torino, Italy.

JANTZEN, MR. IVAR, Nationalbanken, København, Denmark.

JENÖ, DR. ANDREICH, Müegyetem, Budapest, Hungary.
JOFFE, MR. S. A., 515 West 119th Street, New York, N. Y.
JOHN, DR. ERNST, Keilgasse 13/9, Wien III, Austria.
JOHNSTON, MR. V. V., 423 4th Street, Elwood City, Pennsylvania.
JONES, MR. HERBERT E., Cowles Commission For Research in Economics, 301 Mining Exchange Building, Colorado Springs, Colorado.
JORDAN, PROF. CHARLES, Maria-utca 46, Budapest VIII, Hungary.
JUDIK, DR. JÓZSEF, Ruszti ut 10, Budapest, Hungary.
JULIN, MR., A., 81 rue Louis Hap, Bruxelles, Belgium.
JUTILA, MR. K. T., Dept. of Agricultural Economics, Univ. of Helsinki, Helsinki, Finland.

KALDOR, DR. NICHOLAS, London School of Economics, Houghton Street, Aldwych, London, England.
KALECKI, MR. MICHAL, The Rockefeller Foundation, 20 rue de la Baume, Paris VIII, France.
KARRER, DR. S., Director of Research, Consolidated Gas, Electric Light & Power Company, Baltimore, Maryland.
KARVAS, DR. IMRICH A., Kapitulská, Bratislava, Czechoslovakia.
KAUFMANN, PRIVATDOZENT FELIX, Doblinger Haupstrasse, Wien XIX, Austria.
KAUL, DR. BHAGIRATH NATH, 301 List Road, Lahore, India.
KAUTSKY, MR. EDWARD P., 3101 West 44th St., Des Moines, Iowa.
KEILHAU, PROF. WILHELM, Pilestredet 94, Oslo, Norway.
KELLER, DR. HENRY, JR., Dept. of Agricultural Economics, Rutgers University, New Brunswick, New Jersey.
KELLER, PROF. J. O., Assistant to the President, In Charge of Extension, The Pennsylvania State College, State College, Pennsylvania.
KELLOGG, PROF. LESTER, College of Commerce and Administration, Ohio State University, Columbus, Ohio.
KEMÉNY, DR. DEZSO, Szemere-Utca 6, Budapest V, Hungary.
KENÉZ, DR. BÉLA, Tudományegyetem, Budapest, IV, Hungary.
KEYNES, MR. JOHN M., King's College, Cambridge, England.
KING, MR. K. P., Research Department, Central Bank of China, 15 The Bund, Shanghai, China.
KING, PROF. W. I., School of Commerce, New York University, Washington Square East, New York City.
KLEEREKOPER, DR. S., Johannes Verhulststraat 32, Amsterdam Z, The Netherlands.
KLONOV, DR. VLADIMÍR, Institute of Agricultural Economics, Sadová, Praha-Dejvice, Czechoslovakia.
KNAUTH, MR. OSWALD, 366 Fifth Avenue, New York City.
KOLODZIEJCZYK, MR. STANISLAW, Statistical Laboratory, S.G.G.W., Rakowiecka 8, Warszawa, Poland.
KOOPMANS, MR. J. G., Sweelinckplein 29, 's Gravenhage, The Netherlands.
KOOPMANS, MR. T., Netherlands Economic Institute, Pieter de Hoochweg 122, Rotterdam, The Netherlands.
KOVERO, DR. ILM, Professor of Public Finance, University of Helsinki, Helsinki, Finland.
KOZELKA, DR. R. L., University of Minnesota, Minneapolis, Minnesota.
KREPS, DR. T. J., Stanford University, Palo Alto, California.
KROGVIK, MR. SIGURD, Odens gt. 31 B., Oslo, Norway.
KROMPHARDT, PROF. W., Universität-Johannisstr. 9, Muenster-Westf, Germany.

KÜHNE, DR. OTTO, Kurfürstendamm 113, Berlin-Hallensee, Germany.

KULL, DR. WALTER, International Labour Office, Genève, Switzerland.

KURIMURA, DR. YUKICHI, Faculty of Law and Letters, Kyushu Imperial University, Fukuoka, Japan.

KURTZ, DR. EDWIN, 221 River Street, Iowa City, Iowa.

KUZNETS, DR. SIMON, 2 Bent Road, Bowling Green, Media, Pennsylvania.

LANDAU, MR. LUDWIK, Polish Institute for Economic Research, Elektoralna 2, Warszawa, Poland.

LANDRY, M. ADOLPHE, 10 avenue du Square, Paris XVI, France.

LANGE, DR. OSKAR, Rockefeller Foundation, 20 rue de la Baume, Paris VIII, France.

LANGE-NIELSEN, MR. FR., Klingenberggt. 2, Oslo, Norway.

LASORSA, PROF. GIOVANNI, Ministero delle Corporazioni, Roma, Italy.

LATSHAW, DR. V. V., Department of Mathematics, Lehigh University, Bethlehem, Pennsylvania.

LEAVENS, MR. DICKSON H., Cowles Commission for Research in Economics, 301 Mining Exchange Bldg., Colorado Springs, Colo.

LE CORBEILLER, DR. PH., 278 boulevard Raspail, Paris, France.

LEDERER, PROF. EMIL, 3247 Johnson Ave., New York City, N. Y.

LEDERER, DR. WALTER, 545 West 111th Street, New York City.

LEDUC, DR. GASTON, 51 boulevard Beauséjour, Paris XVI, France.

LEGUIZAMON, SR. GUILLERMO E., c/o Directorio Local, Ferro Carril Oeste, Cangallo 564, Buenos Aires, Argentina.

LE NORCY, M. F., French Embassy, 3 Queen's Gate Gardens, London, S.W. 7, England.

LEONTIEF, PROF. WASSILY, 37 Holyoke House, Harvard University, Cambridge, Mass.

LERNER, MR. A. P., London School of Economics, London, W.C.2, England.

LÉVY, PROF. PAUL, 40 rue de la Brasserie, Ixelles-Bruxelles, Belgium.

LIENAU, MR. C. C., Victoria Hotel, Detroit, Michigan.

LIEU, DR. D. K., The China Institute of Economic and Statistical Research, 88 Route Margaret Maresca, (P. O. Box No. 4019), Shanghai, China.

LIMANOWSKI, PROF. ZYGMUNT, 11 rue Poznanska, Warszawa, Poland.

LIMPERG, PROF. TH., JR., Emmalaan 6, Amsterdam Z, The Netherlands.

LINDAHL, PROF. ERIK, Handelshögskolan, Göteborg, Sweden.

LINDER, DR. ARTHUR, Sulgenauweg 20, Bern, Switzerland.

LINDERS, PROF. F. J., Statistics Department, University of Uppsala, Uppsala, Sweden.

LIPINSKI, PROF. EDWARD, Olimpijska 39, Warszawa XII, Poland.

LIVI, PROF. LIVIO, R. Istituto Superiore di Scienze Sociali e Politiche, Via Laura, 48, Firenze, Italy.

LOESCH, DR. AUGUST, Heidenheim (Wuerrt.), Germany.

LOISY, MR. TH., 156 Côte Parée, Toulouse, France.

LORENZ, MR. M. O., 3510 Porter Street N. W., Washington, D. C.

LORIA, PROF. A., Carpa Aparta 26, Torino, Italy.

LOTHERINGTON, MR. JOHN, Statens Pensjonskasse, Oslo, Norway.

LOTKA, DR. ALFRED J., Metropolitan Life Insurance Co., New York City.

LOVEDAY, MR. ALEX, Directeur de la Section financière de la Societé des Nations, Genève, Switzerland.

LOVTCHEVITCH, DR. YOVAN, Vice Gouverneur de la Banque Nationale du Royaume de Yougoslavie, Beograd, Yugoslavia.

LÖWE, DR. ADOLF, 10, Wilbraham Road, Manchester, England.

LUNDBERG, MR. ERIK, Tattby, Saltsjøbaden, Sweden.

LUTFALLA, M. GEORGES, 40, rue de la Voie Verte, Paris XIV, France.

LUZZATTO-FEGIZ, PROF. PIERPAOLO, 8 Via G. Revere, Trieste, Italy.

MACDOUGALL, MR. GEORGE D. A., 6 Lyddon Terrace, Leeds, 2, England.

MACHLUP, DR. FRITZ, Dept. of Economics, University of Buffalo, Buffalo, N. Y.

MCINTYRE, MR. FRANCIS E., Cowles Commission for Research in Economics, 301 Mining Exchange Building, Colorado Springs, Colorado.

MCLEAN, MR. W., c/o University Buildings, Marino, Madras, India.

MADHAVA, PROFESSOR K. B., University of Mysore, Mysore, India.

MADOW, MR. W. G., Fayerweather Hall, Columbia University, New York City.

MAHAIM, PROF. ERNEST, 9 rue du Hêtre, Cointe-Liège, Belgium.

MAKOWER, MISS HELEN, Hatton Garden 56, London, W.C.1, England.

MALKARY, MR. HENRY J., 9 Polo Road, Great Neck, Long Island, N. Y.

MANOILEJO, M. MIHAIL, Professeur d'Économie Politique, l'École Polytechnique, Rue Mussolini 27, Bucuresti, Rumania.

MARGET, PROF. ARTHUR W., School of Business Administration, University of Minnesota, Minneapolis, Minnesota.

MARLIO, M. LOUIS, 23 rue Balzac, Paris VIII, France.

MARN, MR. MATTSON L., Söderhamm, Sweden.

MARSCHAK, DR. JAKOB, All Souls College, Oxford, England.

MARSHALL, DR. L. C., 7007 Rolling Road, Chevy Chase, Md.

MARTIN, DR. FRANZ, Porzellangasse 16, Wien IX, Austria.

MASCI, PROF. GUGLIELMO, Facoltà di Giurisprudenza, R. Università, Città Universitaria, Roma, Italy.

MASNATA, DR. ALBERT, Post Box 4, Lausanne, Switzerland.

MASSÉ, M. P., 9 rue Louis Murat, Paris VIII, France.

MAVERICK, DR. LEWIS A., University of California at Los Angeles, Los Angeles, California.

MAYER, PROF. HANS, Universität Wien, Wien, Austria.

MAYER, DR. JOSEPH, 2852 Ontario Road, N.W., Washington, D. C.

MAYER, MR. LEO KENNETH, 215 Middle Neck Road, Great Neck, Long Island, N. Y.

MEADE, MR. J. E., Hertford College, Oxford, England.

MEANS, MR. GARDINER C., National Resources Committee, Interior Building, Washington, D. C.

MEARS, PROFESSOR ELIOT GRINNELL, 593 Gerone Road, Stanford University, California.

MEDICI, DR. GIUSEPPE, R. Istituto Superiore Agrario di Perugia, Perugia, Italy.

MEDOLAGHI, PROF. PAOLO, Direttore Generale, Istituto Nazionale Fascista della Providenza Sociale, Via Minghetti, N. 17, Roma, Italy.

MELLER, MR. WILLIAM, Ken-u 8, Budapest, Hungary.

MENDERSHAUSEN, DR. HORST, Department of Statistics, University of Geneva, Genève, Switzerland.

MENGARINI, PROF. PUBLIO, 2 Via Monte di Pieta, Torino, Italy.

MENGER, PROF. KARL, Box 206, Notre Dame, Indiana.

MIDUTANI, PROF. K., University of Commerce, Kobe, Japan.

MIGUEL, MR. ANTONIO DE, Orellano 10, Pral. Madrid, Spain.

MILHAU, PROF. JULES, Rue Blanchard, Agde (Hérault), France.

MILLIKAN, MR. MAX, 163 Yale Station, Yale Graduate School, New Haven, Connecticut.

MILLS, PROF. FREDERICK C., Columbia University, New York City.

MILLS, MR. OGDEN L., 15 Broad Street, New York City.

MINOST, MR. ÉMILE, Administrateur-Directeur Générale du Crédit Foncier Egyptien, Cairo, Egypt.

MISES, PROF. DR. LUDWIG, Wollzeile 24, Wien, I, Austria.

MITCHELL, PROF. WESLEY C., 161 W. 12th Street, New York City.

MITNITZKY, DR. MARK, Ermelleki-u 6, Budapest, Hungary.

MOLINA, MR. E. C., Switching Theory Engineer, Bell Telephone Laboratories, 463 West Street, New York, N. Y.

MOLINARI, DR. ALESSANDRO, 15 Via Claudio Monteverdi, Roma, Italy.

MONTGOMERY, MR. J. K., Institut International d'Agriculture, Roma, Italy.

MOORE, PROF. H. L., Cornwall, New York.

MORET, DR. JACQUES, 19 rue de la Convention, Paris XV, France.

MORGENSTERN, DR. OSKAR, Zeitschrift für Nationalökonomie, Schottengasse 4, Wien 1, Austria.

MORINI-COMBY, M. JEAN, 20 rue Clerisseau, Nimes, France.

MORTARA, PROF. GIORGIO, Direzione del Giornale degli Economisti e Rivista di Statistica, Piazzale Sempione 3, Milano 1-44, Italy.

MORTON, DR. J. E., Soběslavská 17, Praha, XII, Czechoslovakia.

MOURRÉ, BARON CHARLES, 5 rue du Boccador, Paris VIII, France.

MOWBRAY, PROF. A. H., 119 South Hall, University of California, Berkeley, California.

MUDGETT, PROF. BRUCE D., University of Minnesota, Minneapolis, Minn.

MYRDAL, PROF. GUNNAR, Norrtullsgatan 2, Stockholm, Sweden.

NAESS, MR. RAGNAR D., Goldman, Sachs & Co., New York City.

NAKAYAMA, PROF. ICHIRO, Commercial University of Tokyo, Tokyo, Japan.

NATHAN, DR. OTTO, 121 W. 11th St., New York City.

NAVRATIL, PROF. DR. ÁKOS, Apponyi tér 1, Budapest IV, Hungary.

NEIFELD, DR. MORRIS R., 883 Brooklyn Avenue, Brooklyn, New York.

NEISSER, DR. HANS, 227 Kenyon Avenue, Swarthmore, Pennsylvania.

NEUBAUER, DR. JULIUS, Iskola utca 13, Budapest II, Hungary.

NEYMAN, DR. J., Dept. of Applied Statistics, University College, London, W. C. 1, England.

NICHOL, MR. A. J., 3412 Otis Street, Mt. Ranier, Maryland.

NIXON, MR. JAMES WILLIAM, International Labour Office, Genève, Switzerland.

NORTON, DR. J. PEASE, 460 Ocean Avenue, West Haven, Conn.

NOYES, MR. C. REINOLD, 53 East 66th Street, New York City.

ODATE, MR. YOSHIO, The Nippon Life Assurance Company, 7, Imabashi Shichome, Osaka, Japan.

OHLIN, PROF. BERTIL, Handelshögskolan, Stockholm, Sweden.

ONICESCU, PROFESSOR OCTAV, Strada Rozeloz 9, Bucuresti, Rumania.

OPIE, DR. REDVERS, Magdalen College, Oxford, England.

OPPEGAARD, MR. K. F., Munkedams Vn. 71, Oslo, Norway.

ORE, PROF. OYSTEIN, Yale University, New Haven, Conn.

PABST, DR. WILLIAM R., Amherst College, Amherst, Massachusetts.

PALMSTROM, HRR. HENRIK, Livsforsikringsselskapet Brage, Oslo, Norway.

PANKRAZ, DR. OTOMAR, Charles University, U Karlova 3, Praha II, Czechoslovakia.

PAPI, PROF. UGO, Via Plinio 7, Roma, Italy.

PARADISO, DR. L. J., 1717 R Street, N.W., Washington, D. C.

PAUER, SENOR PAUL SILICEO, 6a., Naranjo No. 158, Mexico City, Mexico.

PAYNE, DR. CHARLES K., Washington-Square College, New York, N. Y.

PEARSON, MR. EGON S., Senior Lecturer, University of London, University College, Gower Street, London, W.C. 1, England.

PEARSON, PROF. F. A., Cornell University, Ithaca, New York.

PEDERSEN, DR. JORGEN, Institute of Economics and History, Københavns Universitet, København, Denmark.

PEDERSEN, PROF. P. O., Amalievej 1, København, Denmark.

PERLO, MR. VICTOR, Federal Home Loan Bank Board, Washington, D.C.

PERROUX, PROF. FRANÇOIS, Faculté de Droit de l'Université, 106 rue de l'Hôtel de Ville, Lyon, France.

PERSONS, DR. WARREN M., 27 Everett Street, Cambridge, Mass.

PETER, DR. HANS, Freiackerstrasse 53, Tübingen, Germany.

PETERSON, MR. O. STRANGE, Stadion Alle 32, Aarhus, Denmark.

PETRICSKO, PROF. NIKOLAUS, Vigado-tér 1, Penzintezeti Kozpont, Budapest IV, Hungary.

PHILIPPE, PROFESSOR A., Chambre des Députés, Paris, France.

PHILIPS, DR. AUGUST, JR., c/o The Netherlands Bank, Rokin, Amsterdam C, The Netherlands.

PICARD, M. ÉMILE, Secrétaire perpetuel de l'Academie des Sciences, 25 quai Conti, Paris, France.

PIEKALKIEWICZ, M. JEAN, Czerniokόw Miasto Ogród, Goraszewska 4, Warszawa, Poland.

PIETRA, PROF. COMM. GAETANO, Regia Università, Padova, Italy.

PIROU, PROFESSEUR G., Faculté de Droit, Université de Paris, Paris, France.

PIXLEY, DR. HENRY, Wayne University, Detroit, Michigan.

POSE, M. ALFRED, Le Directeur Général, Banque Nationale pour le Commerce, et l'Industrie, Paris, France.

POSTHUMA, MR. S., 10 Acacialaan, Bloemendaal, The Netherlands.

PRIBRAM, PROF. DR. KARL, The Brookings Institution, 722 Jackson Place, Washington, D. C.

PRICE, DR. IRENE, Department of Mathematics, State Teachers' College, Oshkosh, Wisconsin.

PULLERITS, MR. ALBERT, Central Bureau of Statistics, Tallinn, Estonia.

QVALE, MR. PAUL, Solheimsgaten 5B, Oslo, Norway.

RASTOKIN, DR. ING. PETER D., Zemědělskaj ústav Učetnicko-Spravovědny, Praha, XIX, Czechoslovakia.

RAUTENSTRAUCH, DR. WALTER, Department of Industrial Engineering, Columbia University, New York City.

RAWLES, DEAN THOMAS H., Colorado College, Colorado Springs, Colorado.

RAYMOND, PROF. FAIRFIELD E., Mass. Institute of Technology, Cambridge, Mass.

REITSMA, IR. J. M. I., Krabbenbosweg, 120, Hengelo (O), The Netherlands.

RHODES, DR. E. C., London School of Economics, Houghton Street, London, W. C. 2, England.

RICCI, PROF. UMBERTO, Faculty of Law, Giza Branch, Giza, Egypt.

RICHARDSON, MR. BURDICK, Lane Building, Davenport, Iowa.

RICHTER-ALTSCHAFFER, DR. HANS, Bellevuestrasse 8, Berlin, W.9, Germany.

RIDER, PROF. PAUL R., Washington University, St. Louis, Missouri.

RIETZ, PROF. HENRY LEWIS, Physics Building, University of Iowa, Iowa City, Iowa.

RIPLEY, MR. JULIEN A., JR., Oyster Bay, New York.

RIST, PROF. CHARLES, 18 bis rue du Parc de Clagny, Versailles, France.

ROA, MR. EMETERIO, Actuary, Insular Life Assurance Co., Ltd., P. O. Box 128, Manila, Philippine Islands.

ROBBINS, PROF. LIONEL C., London School of Economics, Houghton St., Aldwych, London, England.

ROBINSON, MRS. JOAN, 3 Trumpington Street, Cambridge, England.

ROGERS, PROF. J. HARVEY, Yale University, New Haven, Conn.

ROLL, PROFESSOR ERICH, 29 Hallgate, Cottingham, East Yorkshire, England.

ROMAN, DR. FREDERICK W., 214 Loma Drive, Los Angeles, California.

ROMBOUTS, ING. ALBERT L., Juliana van Stolberglaan 94/96, 's Gravenhage, The Netherlands.

ROMIG, MR. HARRY G., Bell Telephone Laboratories, 463 West Street, New York City.

ROODKOWSKY, MR. NIKITA D., National Arts Club, 15 Gramercy Park, New York City.

ROOS, DR. CHARLES F., Mercer-Allied Corporation, 420 Lexington Avenue, New York City.

ROSENBORG, DR. A., League of Nations, Genève, Switzerland.

ROSSI, PROF. LIONELLO, Via Amalasunta 7, Roma, Italy.

ROY, PROF. RENÉ, 7 rue de la Cité universitaire, Paris XIV, France.

ROYOT, M. GEORGES, Attaché, La Banque des Règlements Internationaux, Basle, Switzerland.

RUEFF, M. JACQUES, 4 rue Pierre Curie, Paris V, France.

RUTSKI, DR. J., Mostowa 3a-20, Wilno, Poland.

SACHS, DR. ALEXANDER, Lehman Corporation, 1-3 South William Street, New York City.

SANKPAL, PROF. G. M., Sydenham College of Commerce, Bombay, India.

SARLE, DR. C. B., U. S. Dept. of Agriculture, Washington, D. C.

SASULY, MR. MAX, Brookings Institution, 722 Jackson Place, Washington, D. C.

SAWKINS, MR. D. T., Industrial Commission of New South Wales, Queen's Square, Sydney, Australia.

SCATES, DR. DOUGLAS, Director of Research, Cincinnati Public Schools, 216 East Ninth Street, Cincinnati, Ohio.

SCHARFF, MR. MAURICE R., One Wall Street, New York, N. Y.

SCHIFF, DR. ERICH, Skodagasse 19, Wien VIII, Austria.

SCHINDLER, DR. A., Schöneiche, Rahnsdorfer Strasse, 50, Post Fichtenau, Germany.

SCHLESINGER, DR. KARL, Alserbachstrasse 16, Wien IX, Austria.

SCHNEIDER, DR. ERICH, Strandparken 23, Aarhus, Denmark.

SCHOENBAUM, DR. EMIL, Professeur a l'Université, Praha, Czechoslovakia.

SCHULTZ, PROF. HENRY, Social Science Building, University of Chicago, Chicago, Illinois.

SCHULTZ, PROF. T. W., Department of Economics, Iowa State College, Ames, Iowa.

SCHUMACH, MR. G. J. D., 74 Outwoods Drive, Loughborough, Leicestershire, England.

SCHUMPETER, PROF. JOSEPH A., 2 Scott St., Cambridge, Mass.

SCHWARZ, ING. FRANKTISĔK, Prĕmyslova 245, Praha-Podoli, Czechoslovakia.

SCOTT, MR. HAROLD CLAUDE, 78 Mayfield Road, Sanderstead, Surrey, England.

SCOVILLE, MR. JOHN W., Chief Statistician, Chrysler Corporation, Detroit, Michigan.

SECRIST, PROF. HORACE, Bureau of Business Research, Northwestern University School of Commerce, McKinlock Campus, Chicago, Illinois.

SELIGMAN, MRS. EUSTACE, 126 East 74th Street, New York City.

SHACKLE, MR. G. L. S., 76 Ebury Street, London, S.W. 1, England.

SHENG, MR. T., National Tariff Commission, Customs Building, P. O. Box 1280, Shanghai, China.

SHEWHART, DR. W. A., Bell Telephone Laboratories, 463 West Street, New York City.

SHIBATA, MR. KEI, Faculty of Economics, Kyoto Imperial University, Kyoto, Japan.

SHOHAT, PROF. J., Mathematics Department, University of Pennsylvania, Philadelphia, Pennsylvania.

SIMMONS, MR. FREDERIC L., G.M-P. Murphy & Co., 37 Threadneedle Street, London, E.C.2, England.

SINDING, DR. THOMAS, Economist, Gabelsgt. 23, Oslo, Norway.

SIPOS, DR. ALEXANDER, Zsigmond-u. 20, Budapest II, Hungary.

SKRZYWAN, MR. WACLAW, Polish Institute for Economic Research, Elektoralna 2, Warszawa, Poland.

SMITH, MR. E. DILLON, 270 Convent Avenue, New York City.

SMITH, MR. H., Lecturer, Department of Commerce, 19 Abercromby Square, University of Liverpool, Liverpool, England.

SMOLENSKY, DR. F. PIETRO, Condirettore delle Assicurazioni Generali, Trieste, Italy.

SMUTNY, DR. PAVEL, State Statistical Office, Bělského, 2, Praha, VII, Czechoslovakia.

SNYDER, MR. CARL, Federal Reserve Bank of New York, New York City.

SOKOLOVE, MR. H., 3435 Brown St., N. W., Washington, D. C.

SPIETHOFF, PROF. ARTHUR, Pappelsdorfer Alle 25a, Bonn, Germany.

SPITZER, MR. LYMAN, JR., Graduate College, Princeton University, Princeton, New Jersey.

SRAFFA, PROF. PIERO, King's College, Cambridge, England.

STACKELBERG, DR. HEINRICH BARON VON, Uhlandstrasse, 15, Lindenthal, Köln, Germany.

STAEHLE, DR. HANS, Chemin Kreig 3, Genève, Switzerland.

STAHL, MR. CAREL, JR., Willemskade 4, Rotterdam, The Netherlands.

STAMP, SIR JOSIAH, Park Hill Road, Shortlands, Kent, England.

STAUDE, DR. HEINRICH, Schnorrstrasse 26, Leipzig W. 31, Germany.

STEFFENSEN, PROF. DR., Ehlersvej 8, Hellerup, Denmark.

STEINDL, DR. JOSEPH, Strassergasse 36, Wien 19, Austria.

STEPHAN, MR. FREDERICK F., American Statistical Association, 722 Woodward Building, Washington, D. C.

STEVENSON, DEAN R. A., School of Business Administration, University of Minnesota, Minneapolis, Minnesota.

STEWART, DR. W. BLAIR, Reed College, Portland, Oregon.

STJERNQVIST, MR. HENRY, Det Statistisk Department, København, Denmark.

STRIGL, DR. RICHARD, Schulgasse 90, Wien XVIII, Austria.

STUCKEN, DR. RUDOLF, Reichskanzlerstr. 71, Altona-Hochkamp, Germany.

SUGIMOTO, PROF. EIITI, Tokyo University of Commerce, Tokyo, Japan.

SUN, DR. C., Bureau of National Resources, 2 San Yuan Hong, Nanking, China.

SURÁNYI-UNGER, DR. THEO, Francis Joseph University, Szeged, Hungary.

SVENNILSON, MR. INGVAR, Stockholms Socialvetenskapliga Institut, Odengatan 61, Stockholm, Sweden.

SWIANIEWICZ, DR. STANISLAW, Docent, Wilno Universytet, Wilno, Poland.
SZTURM DE SZTREM, MR. EDWARD, Wawelska, 78-17, Warszawa, Poland.

TAITEL, MR. MARTIN, 4615 Chestnut Street, Bethesda, Maryland.
TAKATA, PROF. YASUMA, Faculty of Economics, Kyoto Imperial University, Kyoto, Japan.
TAO, DR. L. K., Director, Institute of Social Sciences, Academia Sinica, Nanking, China.
TASSINARI, PROF. GIUSEPPE, Via Lambro 2, Roma, Italy.
TAUSSIG, PROF. F. W., Harvard University, Cambridge, Mass.
THEISS, DR. EDWARD, Horthy Miklos ut. 144, Budapest I, Hungary.
THESEN, MR. GUDBRAND, Cond. Actuar, Idun, Kirkegaten 24, Oslo, Norway.
THIRRING, DR. LOUIS, Zsigmond-utca 1, Budapest, Hungary.
THOMAS, MR. CARL, 43 Holyoke House, Cambridge, Mass.
THOMPSON, PROF. C. D., Holland Hall, Allahabad University, Allahabad, India.
THOMPSON, DR. JOHN MURRAY, 224 Giannini Hall, University of California, Berkeley, California.
TIMOSHENKO, DR. V. P., Food Research Institute, Stanford University, California.
TIMPE, PROFESSOR, Dianastrasse 21, Berlin-Schlachtensee, Germany.
TINBERGEN, DR. J., 12 avenue du Mervelet, Genève, Switzerland.
TINTNER, PROF. GERHARD, Dept. of Economics, Iowa State College, Ames, Iowa.
TOBATA, PROF. SEIICHI, Faculty of Agriculture, Tokyo Imperial University, Tokyo, Japan.
TOBILCOCK, MR. D. H., c/o Govt. Statistical Bureau, P. O. Box 791, Salisbury, S. Rhodesia, S. Africa.
TOLLEY, PROF. HOWARD R., Room 3095, South Building, Dept. of Agriculture, Washington, D. C.
TOMLINSON, DR. FREDERICK ROTHMAN, Division of Economics and Markets, Union Buildings, Pretoria, So. Africa.
TOUT, DR. HERBERT, 10 Royal Park, Clifton, Bristol 8, England.
TREVOUX, DR. F., Faculté de Droit de Grenoble, Grenoble, France.
TRUKSA, DR. LADISLAV, Praha XI-1800, Czechoslovakia.
TSHA, MR. T. Y., 575 Rue Amiral Bayle, Shanghai, China.
TUTTLE, DR. CHARLES R. M., 1831 Jefferson Place, N.W., Washington, D. C.

UGARTE, MR. SALVADOR, Calle V. Carranza No. 44, Mexico City, Mexico.
UGGÉ, PROF. ALBINO, Via Giuseppe Mazzini N. 25, Milano VI-H, Italy.

VALK, DR. WILLIAM L., Weissenbruchstraat, 398, 's Gravenhage, The Netherlands.
VANDELLOS, DR. JOSÉ A., Rosellón 253, 5°, 2ª, Barcelona, Spain.
VAN DER MEER, MR. BARELD, Nassauplein 22, 's Gravenhage, The Netherlands.
VAN DER WIJK, DEN HEER, Corn. Krusemanstr. 27^{II}, Amsterdam Z, The Netherlands.
VAN DORP, DR. E. C., Villa Jamga, Bloemendaal, The Netherlands.
VAN GENECHTEN, DEN HEER G., Willemsplantsoen 10, Utrecht, The Netherlands.
VAN PRAAG, DR. IR. M. M., 49 Nassaulaan, Schiedam, The Netherlands.
VAN SICKLE, DR. JOHN V., c/o Rockefeller Foundation, 49 West 49th Street, New York.
VAN WINKLE, PROF. E. H., Rensselaer Polytechnic Institute, Troy, New York.
VARGA, DR. STEPHEN, Director, Hungarian Institute for Economic Research, 8, Alkotmány-utca, Budapest, Hungary.
VERGARA, PROF. ROBERTO, Direccion General de Estadistica, Teatinos 307, Santiago, Chile.

VERRYN-STUART, PROF. G. M., Emmalaan 6, Utrecht, The Netherlands.
VIAL, DR. EDMUND E., Bureau of Agricultural Economics, Washington, D. C.
VIDAS, DR. PIERO, Via S. Basilio 14, Roma, Italy.
VINCI, PROF. FELICE, Viale Aldini 3, Bologna, Italy.
VINOGRADOFF, MR. D. I., Westinghouse Electric & Manufacturing Co., East Pittsburgh, Pennsylvania.
VLEUGELS, PROF. DR. WILHELM, Argelanderstrasse 12, Bonn am Rhein, Germany.
VOLTERRA, SENATOR VITO, 17 Via in Lucina, Roma, Italy.
VON BECKERATH, PROF. ERWIN, Universität Köln, Köln, Germany.
VON BECKERATH, DR. HERBERT, 203 Battle Lane, Chapel Hill, North Carolina.
VON MERING, BARON OTTO, 18, Villa Anna, Davos-Platz, Switzerland.
VON NEUMANN, PROF. JOHN, 65 Stockton Street, Princeton, New Jersey.
VON SZELISKI, MR. VICTOR S., 3318 Quesada St., Washington, D.C.

WAFFENSCHMIDT, PROF. W., Bismarckstrasse 59, Karlsruhe, Germany.
WAGEMANN, PROF. DR. E., Fasanenstrasse 6, Berlin-Charlottenburg 2, Germany.
WAKAR, DR. ALEKSY, ul. Mokotowska 57, Warszawa, Poland.
WALKER, PROF. HELEN M., Teachers College, Columbia University, New York, N. Y.
WARREN, PROF. G. F., Cornell University, Ithaca, New York.
WAUGH, DR. FREDERICK V., U. S. Bureau of Agricultural Economics, Washington, D. C.
WEAVER, MR. DONALD, 1332 Vermont Ave., N. W., Washington, D. C.
WEDERVANG, PROF. INGVAR, Anton Schjöthsgt. 14, Oslo, Norway.
WEIDA, PROF. FRANK M., The George Washington University, Washington, D. C.
WELCH, MR. B. L., 10 Holmlands Park South, Sunderland, England.
WELLS, ACTING PRESIDENT HERMAN B., Indiana University, Bloomington, Indiana.
WENZLICK, MR. ROY, Real Estate Analysts, Inc., St. Louis, Missouri.
WESTERFIELD, PROF. R. B., Yale University, New Haven, Conn.
WHITMAN, DR. ROSWELL H., Office of the Controller, R. H. Macy & Co., New York City.
WICKSELL, PROF. SVEN DAG, University of Lund, Lund, Sweden.
WILSON, PROF. E. B., School of Public Health, Harvard University, Shattuck Street, Boston 17, Mass.
WILSON, MISS ELIZABETH, 1 Waterhouse Street, Cambridge, Massachusetts.
WINKLER, PROF. DR. WILHELM, Universitätsinstitut für Statistik, Neue Burg, Heldenplatz, Wien I, Austria.
WISNIEWSKI, PROF. JAN, Rakowiecka 6, Warszawa, Poland.
WOLD, DR. HERMAN O. A., University, Stockholm, Sweden.
WOLF, MR. A. F., Hills Brothers Company, 110 Washington Street, New York City
WOLFF, M. ROBERT, 130 avenue Victor Hugo, Paris, France.
WOODBURY, DR. ROBERT M., c/o International Labour Office, Genève, Switzerland.
WORKING, PROF. E. J., Department of Agricultural Economics, University of Illinois, Urbana, Illinois.
WORKING, PROF. HOLBROOK, Food Research Institute, Stanford University, California.
WOYTINSKY, MR. WLADIMIR, 4943 Butterworth Place, N. W., Washington, D. C.
WU, MR. TAYEH, Nankai Institute of Economics, Tientsin, China.

YNTEMA, PROF. THEODORE O., University of Chicago, Faculty Exchange Box 243, Chicago, Illinois.

ZAGOROFF, DR. SLAVCHO, Parchevich 32, Sofia, Bulgaria.

ZAGORSKI, MR. JOZEF, Polish Institute for Economic Research, Elektoralna 2, Warszawa, Poland.

ZASSENHAUS, DR. HERBERT, 11 Wendell Street, Cambridge, Massachusetts.

ZAWADZKI, PROF. WLADYSLAW, Alyna na Skarpie 3, Rakowieka 6, Warszawa, Poland.

ZAYCOFF, DR. RASHCO, Statistical Institute for Economic Research, State University, Shipka 6, Sofia IV, Bulgaria.

ZEUTHEN, DR. F., Hastrups, Have 1, København V., Denmark.

ZORNER, PROF. DR. HANS, Schillerstr. 12a, Berlin-Lichterfelde-Ost, Germany.

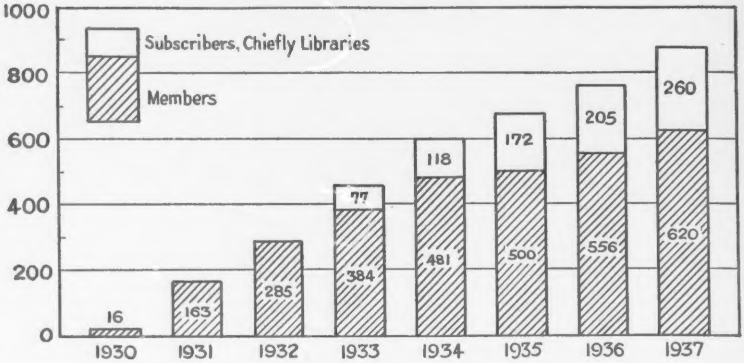
ZUJOVIC, PROF. MILAN J., Prof. à la Faculté de Droit, Kralja Milana 42, Beograd, Yugoslavia.

ZWEIG, DR. FERDYNAND, University of Kraków, Lobzowska 4, Kraków, Poland.

SUMMARY

Members, October 1, 1936		557
New Members Added		83
		640
Less: Deaths	4	
Resignations	16	20
Members, October 1, 1937		620

Growth and Distribution of Membership in the Econometric Society and of Subscribers to Econometrica



Geographic Distribution of Members and Subscribers

	Members	Subscribers	Total		Members	Subscribers	Total
United States	216	108	324	Bulgaria	4	1	5
England	46	21	67	Canada	2	3	5
Italy	45	22	67	Egypt	4	1	5
France	52	9	61	Yugoslavia	4	1	5
Germany	22	9	31	Finland	3	1	4
Netherlands	26	5	31	Scotland	2	2	4
Japan	10	13	23	South Africa	3	1	4
Poland	17	6	23	U.S.S.R		4	4
Hungary	21	1	22	Chile	2	1	3
China	12	8	20	Dutch E. Indies		2	2
Switzerland	18	2	20	Greece		2	2
Sweden	15	4	19	Mexico	2		2
Norway	15	3	18	Algeria	1		1
India	9	7	16	Brazil		1	1
Denmark	14	1	15	El Salvador	1		1
Czechoslovakia	13	1	14	Estonia	1		1
Belgium	12	1	13	Lithuania		1	1
Austria	11		11	Philippines	1		1
Australia	4	6	10	Puerto Rico		1	1
Argentina	3	4	7	Tasmania		1	1
Rumania	6	1	7	Uruguay		1	1
Spain	3	4	7				
				Total	620	260	880

C15